SYLLABUS – Fall 2015

CSCI E 59 Database Management

Course Prerequisite CSCI E-10a, CSCI E-20, or the equivalent.

Instructor(s)

Dr. Ramon A. Mata-Toledo, Instructor
Email: rmatatoledo@fas.harvard.edu (preferable) or rmatatoledo@gmail.com

Professor: Larry L. Adams II, Teaching Assistant
Email: professorlarryadams@gmail.com

Contacting the instructors

Please email Dr. Mata-Toledo and/or Professor Adams with any personal or private issues that are affecting your class experience or with any information you do not wish to share with the rest of the class. Example: death in the family, severe illness, grading issues, etc.

Please do NOT email Dr. Mata-Toledo and/or Professor Adams with questions about the week’s material. It is better to ask these questions in the discussion forum for the class, where others can benefit from the answer as well. (See “How Class Works” below.)

Meeting Time:

Lectures and QA Session: Thursday from 7:40 p.m. until 9:40 p.m.
Dr. Mata-Toledo

Labs and QA Session: Saturday(s) at 10:00 a.m. until 12:00 p.m.
Professor Larry Adams hands on approach to SQL

Background and course content

The main objective of this course is to understand the basic concepts necessary to create a functional database free of data anomalies. Although a strong mathematical background may be necessary to handle advance concepts multivalve dependencies and representation theory, most students can conceptualize, design, and implement a database with no more than high school mathematics and good common sense. Concepts such as representation theory are beyond the scope of this course.

The course will use a mixture of lecture notes, readings, labs, and homework.

How the class works

1. The entire class is taught online through canvas and blackboard collaboration which means you have flexibility in attending the class. It is highly recommended that you attend the lectures and labs in order to interact with your
Professor, TA, and fellow students. But we understand that most of you are working students and at Harvard University it is our goal to work with my diverse schedules so you may continue your journey in higher learning.

2. Each week we list our assignments, homework, labs, and projects for you to stay focus. There is a weekly schedule so you will not be shocked by a upcoming due date. IT IS HIGHLY RECOMMENDED that you plan your schedule out.

3. We have students from around the world in this class so we try to make our lectures and labs precise and clear. If you have any question PLEASE feel free to ask the question. Our goal is to help you to understand and implement the complexities of database design in the business world with real hands on learning exercises.

4. There is a discussion forum where you can interact with your fellow students. We will answer your questions and post discussion topic to peak your interest in trends that are going on in the IT communities.

5. Each class builds on your knowledge so it is imperative that you keep up with the work. We have tried to spread out the assignments, lectures, labs, and final project to ensure you have enough time to complete everything without stress.

Course Objectives:

The object of this course is to teach students the general concepts of relational databases and how to design a database that is anomaly free. Students will learn to design, create, populate, and query a database by working with the Oracle ™ database engine and the SQL language. Students will also learn basic database administration skills such as creating users, granting/revoking privileges individually or collectively to several users through the use of roles. The main objectives of this course are:

- To emphasize the features of a Relational Database, in particular, the most widely used model, the relational model, and the use of the Oracle ™. Know the basic components of a relational database management system (RDBMS) and how these parts interact during the normal activities of a database engine.

- To construct a conceptual model (E/R diagram) and a physical model (Relational Design) from a general data description that runs in an Oracle ™ engine. Create and populate the database and impose the necessary constraints to satisfy data integrity and operational requirements.

- To normalize a database from 1NF to BCNF and the necessity of this process to avoid data anomalies by identifying partial and transitive functional dependencies by the correct application of Armstrong’s inference axioms.

- To write queries in SQL that implement the theoretical relational database operators such as selection, projection, joins. Learn how to impose and enforce constraints on the data to satisfy the requirements of some of the normal forms.

- To write queries in SQL that uses features such as group function (AVG, MIN, MAX, COUNT), individual functions (arithmetic functions, character, conversion functions), processing of date and time information using functions such as TO_DATE, SYSDATE pseudo columns.

- To write basic reports using the BREAK and COMPUTE commands of SQL Plus™.

- To write complex queries (subqueries, correlated queries, update tables, inserting/deleting rows from tables using subqueries) and hiding data through the use of Views.

- To be able to accomplish basic administrative tasks such as create/delete users, grant/revoke privileges to users individually or through roles.

- Managing a successful semester-long project from initial description through design, creation, and implementation.

About the textbooks

In this course we will make use of two textbooks an Introduction to Database Systems by C. J. Date and Fundamentals of SQL Programming by R. A. Mata-Toledo. These two books and some additional references are shown below. Our primary textbook is “An Introduction to Database Systems” by C. J. Date. This textbook presents the rationale and overview of the
relational model; the second book shows you how to implement these concepts in Oracle™. One word about the second book is that it is only “fundamentals of SQL” and even though it was written for Oracle™ 8, all the commands of SQL work in all versions including the latest Oracle incarnation, 12g. Some of the most advanced and controversial commands introduced lately by Oracle™ are not covered here. However, you should be able to query any Oracle database with the confidence that the instructions work the same in any Oracle version. This second book also offers a “practice database” where you can try the different command and how they are applied. This practice database is similar to the one used in the training programs of the Oracle™ Corporation. This database and its ER diagrams are used with expressed permission of the Oracle™ Corporation.

Textbooks


This book provides all the theoretical background that is needed for the course. It also provides an overview of all “practical issues of databases particularly about the relational data model and some of its characteristics. The material that is critical to the course is that of Part III (Database Design).

Should you find the textbook Introduction to Database Systems 8th Edition somewhat expensive you can still use any of the previous editions 5th through 7th. The new edition includes topics that are beyond the scope of this class.


This book is a “how to do book” on Oracle™ and complements the theoretical aspects that Introduction to Database Systems by C.J. Date presents at a “higher level”. Although the Oracle engine used in the books is Personal Oracle 8, however, all commands should work the same in all Oracle ™ versions including Oracle ™ 12g.

Additional Recommended Readings

The following books can be used as reference sources. Although the latest editions are shown here, for purposes of the material covered in this class any of the previous editions can be used.


The student is also encouraged to use any resource available in the Internet regarding the content of this course including the material available at the Oracle™ Corporation website. Both the instructor and the TA will indicate some appropriate websites that will help you complement the material covered in this course.

Graded assignments
Grades in this course will be determined according to the deliverables shown below (weights are shown in parenthesis) and an active participation in the course:

- Homework (40%)
- Participation (20%)
- Final Project (40%)

Graduate students will have additional work or additional requirements to complete for each assignment.

**Grading Matrix and Feedback**

Our goal is to have all work graded within 3 days. However, we do have a late policy that we deduct 10 points per day for 3 days. If you do not turn in your assignment, project, or homework after three (3) days of the due date, your grade becomes a zero. This means you might not receive your grade for 6 days total. It is important for us to provide you timely feedback so you can understand the material as they build on each other. If you do not receive a grade within a week you are to email Professor Adams (professorlarryadams@gmail.com).

**Homework assignments**

The assignments for this course can be divided into three main groups: homework, a project, and individual participation (labs). The homework will allow students to show their mastery of particular SQL commands. The project will allow the students to apply all the concepts of the relation data model in the creation of a database. This professor will provide you with a “general description” of a database that you will design.

It is strongly recommended you complete the assignments. You will find that if you wait for the end of a unit to build the final project, you will have gaps in your understanding if you simply listen to lectures without practicing the skills that are taught.

**Grading procedures**

Projects are due on the dates specified in the course outline by 11:59 PM Eastern time that day. It is expected that you will turn the assignment in on time.

**Turning in assignments**

Your assignment will generally be in the form of a SQL script. Scripts should be text only with the .sql extension. You will need to label your assignments with your first initial, last name, and the name of the assignment. Example ladams_homework1 and zip the file to upload in canvas (ladams_homework1.zip).
Academic Honesty
Once again, the purpose of this course is to teach you the fundamentals of Database Design with Structured Query Language by Oracle and MySQL. This means that copying the files of other students and submitting them as your own work is not only a violation of school academic policies but a real shortchange to your ability to learn what you need to learn. You must write your own code (SQL) or use APA standard to cite your sources in graded assignments.

We fully expect that there will be no issues in this area. We ask that you help keep us out of situations where we might need to take action regarding plagiarism, cheating and other acts of poor academic integrity.

About SQL
This query language is the most common language to interact directly with a database. Learning SQL is like learning a foreign language. At the beginning we may feel overwhelmed but in reality is fairly simple to use once you understand and get familiar with its syntax. Although there are “new features” to SQL, in this course we will be using the “basic commands” that allow us to interact with any database. In this particular course we will be using the Oracle™ flavor of SQL. Yes, there are some other “flavors” of SQL since each database manufacturer tries to make its own product unique, however, think of these flavors as variations of common language. As I mentioned before, learning SQL is learning a foreign language. If you are learning English, you can learn it in England, or is some of the Caribbean Countries, Ireland, South Africa, or Australia. After a while you will know how to understand, read, and write English. Yes, in every country where you study English you will learn different accents, slangs, expressions, and colloquialisms that you can translate, should you have the need, into the other English flavors. It is the same with SQL, once you learn what an instruction does in Oracle SQL, you just need to know how to say it “in the other SQL flavor” by looking at the “new syntax”, however, the result of your queries should be the same.

Project:
The student will design, create, and populate a working database from description requirements provided by the instructor. Students will work in groups generally of two or three persons maximum. It is the responsibility of the students to connect with each other and form their own groups. It is expected that each individual will learn all aspects of the design and creation of the database. However, the project will be presented as a single group. The idea of the project is to take the student not only through the process of creating a database but, more important, of realizing that the design of a database must take into account the minimization of redundancy and the avoidance of anomalies. Avoidance of anomalies is generally sacrificed in the altar of speed and managed in an “ad hoc” manner as “they are encountered”. This is a very common error because what we want is “fast results” not good design. As part of project the student will include the original description of the project, the corresponding ER diagrams, Table Instance Charts, tables, constraints imposed on the data and the corresponding scripts. It is also expected that the student illustrate the use of different commands as they are applied to the database being implemented. The professor will provide a particular set of questions about the database and the students will provide the necessary queries to answer these questions.

Student Responsibilities
Students who enroll in this course are expected to meet the following requirements:

• To be aware of all announcements and to meet deadlines.
• To participate in the discussions board and make substantial contributions
• To maintain a respectful environment that allows all students to achieve the course objectives.
• To abide by university policies and Honor Codes of Conduct

Direct communication with the instructor is encouraged and expected. Learning is “team effort” between the students and the instructor. This instructor welcome all sorts of questions regarding the material covered in the course. If for some reason I do not know the answer immediately, I will research it and give you an answer as soon as I can.

Discussions: The professor will write scenarios so you can comment, critique, rip apart based on your comprehension of the material. The objective of these discussions is to manifest your understanding of the subject. Participation is what is important. If you agree or disagree with a particular statement or point of view, do not worry, that is part of the game of learning databases. The professor will intervene if the discussion is not on track or to clarify what is not clear.

How to contact your professor or Teaching Assistant
It is my objective to teach you how to work with databases. Although you may read books and see examples it is not until you get our hands “dirty” that you start learning the idiosyncrasies of a particular database management system. They all work on the same principle but you need to learn to how to get it work for you. For this reason, feel free to contact this instructor via email at rmatatoledo@fas.harvard.edu or through Canvas. Your TA can be reached through Canvas or via email at professorlarryadams@gmail.com. We will try to respond to all your emails as soon as we can. Dr. Mata-Toledo generally teaches in the mornings from 8:00 a.m. until around 1:00 p.m. So expect answers to your queries after he finishes his classroom lectures. Ask for amplifications of any topic which is not clear to you; I appreciate if you want to facilitate my job by not asking questions but you need to make sure that you learn the material and learn it well. No question is silly; ask what you do not understand. My job is to make sure that you “get it”. If for some reason I do not know the answer to one of your questions immediately, I will research it and provide you an answer. Make sure that in every email you write Harvard Extension in the Subject line followed by whatever topic you are interested in. For example, if it is a question, write Harvard Extension (Question about the normalization process). All your questions will be stored in a folder by the name of Harvard Extension Databases- Fall 2015. As I build a library of questions, I will be able to provide future students with a FAQ section. Working with database is fun, although sometimes to get answers to particular queries may be frustrating, but remember, the journey is generally more fun than just arriving at the destination.

Tentative Schedule

The tentative schedule of the course only shows 14 weeks of activities, this is so because this will allow us to dedicate the necessary time to those topics that are more difficult to the student or give the student some degree of flexibility due to their schedule. Unless something catastrophic happens, God willing, we will try stick to the schedule.

Week No. 1 (September 1, 3, & 5)

What is the benefit of using a database versus a shared file system? What is Data models and the relational database system? Data independence versus data-dependent data and how a database addresses these two issues. The Three-level Architecture and why it is necessary. What are the characteristics of each of these levels and the role of the database administrator in establishing the separation of these levels? What is database management systems, its components and how they work together.

Activities for this week: Read Chapter 1 and 2 of Introduction to Database Systems. Make sure that you understand the definitions at the beginning of the exercise section.

Week No. 2-3

The Entity Relationship Model and the Oracle conventions we will use to express relationships between Entities. ER diagrams, resolution of M:N relationships, and Table Instance Charts (TICs). Translations of TICs into relational tables. Activities for this week: Read Chapter 1 of Fundamentals of SQL and the article Strategies for Determining a Design View: A Preamble to DBMS Modeling.

Graded exercise!

Lab 1 – install and configure the following software – MAMP or WAMP, Oracle Workbench, and Visio.
Due on September 13th!

Homework 1 - Start creating the ER diagram of the database you intend to build. Identify the main entities of the database according to the Oracle conventions.
Due on September 20th.

Week No. 4

Introduction to SQL and relational database concepts. Relations and attributes. Candidate and primary keys. Foreign keys and why they are necessary. Introduction to relational operators and how they are applied. Learning how to interact with Oracle through the SQL*Plus. Creating and deleting tables.

Activities for this week: Read Chapter 1 of Fundamentals of SQL. Read Chapter 6 of Introduction to Database Systems.

Week No. 5

Constraints imposed in a database. Updating and deleting rows in a table using the UPDATE TABLE, DELETE TABLE, and the DROP TABLE command with and without constraints. Implementation of the Selection and Projection operators. Ordering the results of a table according to a given attribute in ascending or descending orders.
Activities for this week: Read Chapter 1 of Fundamentals of SQL and Sections 2.1 through 2.3. Make sure that you can follow the commands of the Solved Problems Section. Read Chapter 7 of Introduction to Database Systems.

Graded exercise!
Homework No. 2 TBA

Week No. 6
Normalization of a database. First, second and third normal forms. How to detect anomalies and use of the Armstrong’s axioms for determining functional dependencies. Importance of normalizing a database and the types of anomalies that may be encountered in First, Second, and Third Normal Forms. How to recognize, prevent, and how to get rid of anomalies in these forms.

Activities for this week: Read Chapters 11 and 12 of Introduction to Databases until 3NF. Read handouts on functional dependencies and the detection of anomalies through the use of Armstrong’s Axioms. Read handout on the normalization process.

Graded exercise!
Lab 2 – Create an ERD with Toad software TBA

Week No. 7
Continuation of the normalization process. BCNF form and Dependency preservation. Algorithms to ensure dependency preservation. The Join operator and its different types. Advantages and disadvantages of higher normal forms from an operational point of view.

Graded exercise!
Homework No. 3 TBA

Activities for this week: Read Boyce/Codd Normal in Chapter12 of Introduction to Databases. Read handout on the normalization process.

Week No. 8
Boolean Operators in SQL and pattern matching using the LIKE clause, % and underscore characters. Arithmetic Operations and use of built-in functions in SQL.
Activities for this week: Read Chapter 3 and 4 of Fundamentals of SQL.

Graded exercise!
Homework No. 4 TBA

Week No. 9
Continuation of Boolean operators if necessary and introduction to Group functions using the Group by clause and additional built in functions. Processing dates and time and basic arithmetic with dates. Formatting of dates and times. Homework No. 6.
Activities for this week: Read Chapter 5 and 6 of Fundamentals of SQL.
Project tables, constraints and populations are due at the end of this week. This assumes that the database has been normalized at least to BCNF form.

Graded exercise!
Homework No. 5 TBA

Week No. 10
Individual project review and normalization issues that need to be resolved. Checking of referential integrities constraints and preservation of data. Activities for this week:

Graded exercise!
Homework No. 6 TBA
Week No. 11
Introduction to complex queries and set operators. Subqueries and correlated queries.

Activities for this week: Read Chapter 7 of Fundamentals of SQL. Students and Professor will review the status of the project at this time and address any concern about its design or implementation.

Graded exercise!
Homework No. 7 TBA

Week No. 12

Activities for this week: Read Chapter 16 and 16 of Introduction to Databases. Read handout on concurrency.

Week No. 13
Security issues using SQL. Authentication, creation and assignment of privileges to users. Views as a mechanism to hiding data from users. Activities for this week:

Graded exercise!
Homework No. 8 TBA

Week No. 13

Activities for this week: Read Appendix E of Fundamentals of SQL. Read handout on reports.

Week No. 14
Human aspects and performance design of databases. Ten commandments of a good design. The why and how’s. Activities for this week: read handouts.

Week No. 15 and 16
Catching up and presentation of individual projects.

At the end of the semester, I expect that you have a good understanding of the issues involved in designing a good database and some of the trade-off that you need to make to balance the internal structure of the database and performance issues. I also hope that you have acquired a good understanding of SQL, its advantages and limitations. Finally, it is my hope that you have enjoyed the course and found worthy the time and effort that you invested. At any time, feel free to let me know how the course can be improved or adapted to your learning style. Be well and enjoy life.