**John Jay College**

**Computer Science Department**

**Course Syllabus**

**CSCI 362: Database and Data Mining - Fall 2020**

|  |  |  |  |
| --- | --- | --- | --- |
| **Instructor** | **Room** | **Meeting time** | **E-Mail** |
| Sanjay Ramchandani | New Bldg 6.61 | MW 4:30 -5:45 pm | sramchandani@jjay.cuny.edu |
| Office Hours: | 10 minutes after class or by appointment | | |

# Course Description

In this course, students will understand the fundamental principles of database management systems (DBMS), and learn different data models including relational, network, object-oriented data model. Through projects, they will gain valuable skills on how to design scalable databases using entity-relationship diagrams. They will learn how to extract useful information from the database by discovering patterns within the underlying data, and learn how to protect data. Finally, they will develop database analyst skills using SQL, and use this to implement several real-life database applications. Students will gain hands-on experience with a modern DBMS platform by completing several exercises in the computer lab. Upon successful completion of this course, students will be able to design and organize a database to manage their application data, and be able to design queries to extract useful information from a database. We will also discuss data mining techniques and algorithms.

# Prerequisite

All students should have completed ENG 201, CSCI 373 or MAT 373 with grade of C or better. If you do not satisfy this requirement, please see me immediately.

# Course Objectives

After completing the course, the students should be able to participate in the design and development of a database. Students are expected to:

* Understand the business requirements and build the conceptual model of a database using Entity Relationship Model.
* Build the relational model of the database by mapping the conceptual model into the set of relations.
* Build the physical model of the database using SQL.
* Be able to develop a DB structure for a specific system/problem.
* Configure a commodity DBMS for secure access.
* Demonstrate basic SQL proficiency for table creation, data insertion and data query
* Describe DBMS access controls and privilege levels and apply them to a simple database
* Know about the usage of Database in Applications.

# Textbooks

* **Modern Database Management** by Jeffrey A Hoffer, V. Ramesh, Heikki Topi, Pearson Education, 13th **edition**
* Introduction to **Data Mining**, by Pang-Ning Tan, Michael Steinbach, Anuj Karpatne , Vipin Kumar , the **second edition** , with ISBN-13 9780073523323

# Software

All programming in this course requires SQL. In Computer Science Department we use the open source **MySQL latest community edition y**ou can install it on your own computer, available as a free download called MySQL. We will also do some mongo db development.

# Grading and homework

Course grades are based on homework, two exams, and a term project.

Assignments 30%

Project 15%

Midterm exam 25%

Final exam 30%

Homework problems are due at the time specified in the assignment description; you can submit it electronically or hand it in class. All homework should be individual work. A penalty of late homework and projects will be applied as follows:

* 0 to 72 hours late: 25% penalty
* 72 to 144 hours late: 50% penalty
* More than 144 hours late: no credit will be given.

# Term project

Term project will give you hands-on experience in the design and implementation of a database. You will work in a team to develop a database application using MySQL or SQL Plus. Specified in a separate document. Identify a team of fellow students who can meet together regularly. **The ideal team size is 3-4**. As we cover material in class, you will apply it to your project. Periodically you will be asked to submit your work for review. After the final projects are submitted, your project will be tested by your classmates Your project must run in SQL.

The project is due and graded in five steps:

**Step 1**: identify appropriate entities, relationships, attributes, domains, and keys as well as all specializations and generalizations.

**Step 2**: produce the associated ER diagram.

**Step 3**: map the ER diagram, and show how you normalized them if necessary, list integrity constraints

**Step 4**: define the database in MySQL and insert at least 10 tuples into every table.

**Step 5**: implement transactions and submit the full project, including an essay on the project.

# Course outline

Unless specified, the referenced chapters are from the textbook **Modern Database Management and the Data Mining concepts are from Introduction to Data Mining.**

|  |  |  |
| --- | --- | --- |
| ***Date*** | ***Topic*** | ***Source*** |
| 8/26 | Context of Database Management | Part 1 |
| 8/31 | Context of Database Management | Part 1 |
| 9/2 | Data models and Database Design | Part 2 |
| 9/9 | The ER model | Part 2 |
| 9/14 | The ER model | Part 2 |
| 9/16 | The ER model | Part 2 |
| 9/21 | SQL introduction | Part 2 |
| 9/23 | No class scheduled |  |
| 9/29 | SQL continued | Part 2 |
| 9/30 | SQL continued | Part 2 |
| 10/5 | Normalization | Part 2 |
| 10/7 | More normalization | Part 2 |
| 10/12 | College is closed – No classes |  |
| 10/14 | SQL meets in the lab and midterm discussion | Part 2 |
| 10/19 | **Midterm** |  |
| 10/21 | Datawarehousing and DI | Part 3 |
| 10/26 | Datawarehousing and DI | Part 3 |
| 10/28 | Big Data Technologies | Part 3 |
| 10/30 | Big Data Technologies | Part 3 |
| 11/2 | Analytics and its implications | Part 3 |
| 11/4 | Analytics and its implications | Part 3 |
| 11/9 | Distributed Databases | Part 3 |
| 11/11 | NoSQL concepts and use cases | NA |
| 11/16 | NoSQL concepts and use cases, mongo db as an example. | NA |
| 11/18 | NoSQL continued | NA |
| 11/23 | Introduction to database security | NA |
| 11/30 | Database vulnerabilities (SQL injection) | NA |
| 12/2 | Data Mining introduction | Chap1 to 4 (DM) |
| 12/7 | Data Mining concepts | Chap1 to 4 (DM) |
| 12/9 | Data Mining algorithms | Chap1 to 4 (DM) |
| 12/11 | **Project Evaluation and Final Examination** |  |

**Academic dishonesty policy:**

Academic dishonesty is unacceptable and will not be tolerated. Academic sanctions in this class will range from an F on the assignment/exam to an F in this course. A report of suspected academic dishonesty will be sent to the Office of the Dean of Students. Additional information and definitions can be found at

[**http://www.hunter.cuny.edu/studentservices/advising/policiessub/proceduresforreporting-academic-integrity-violations**](http://www.hunter.cuny.edu/studentservices/advising/policies-sub/procedures-for-reporting-academic-integrity-violations)