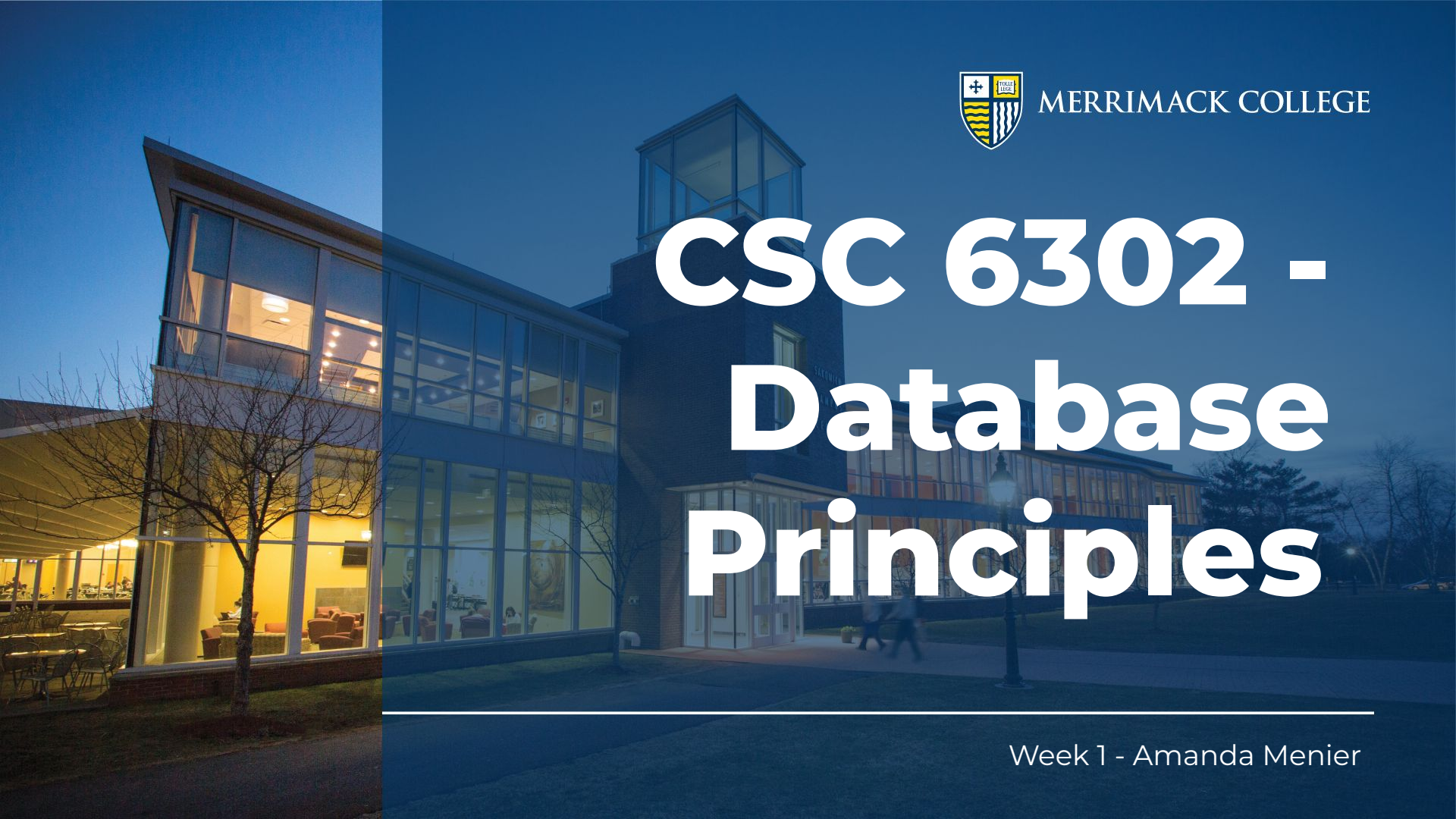


MERRIMACK COLLEGE

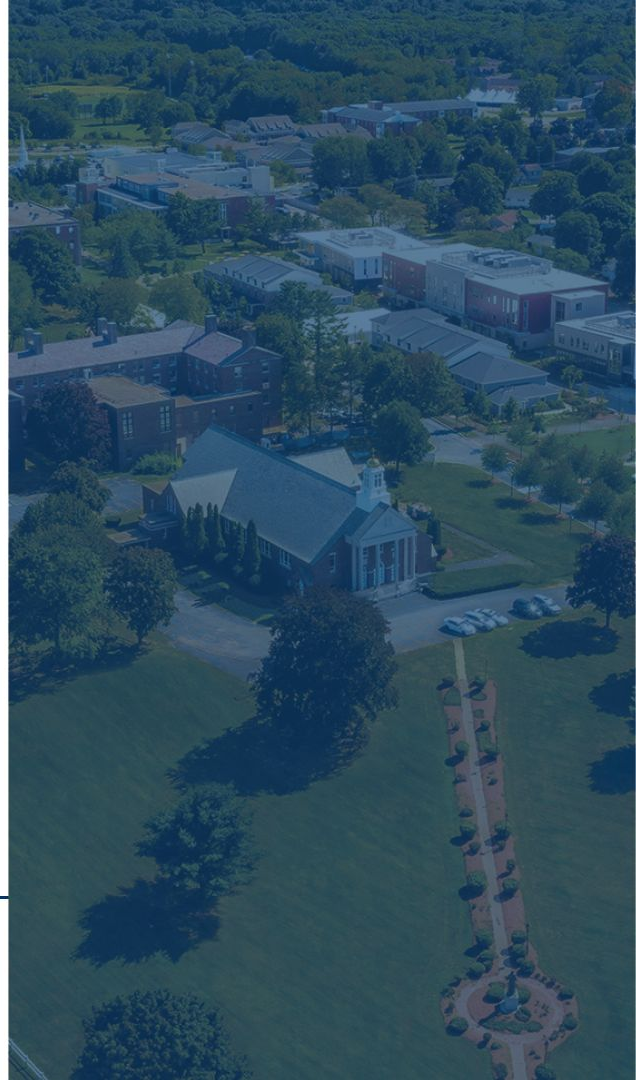
CSC 6302 - Database Principles

Week 1 - Amanda Menier

Agenda

Discussion Topics:

- Course Overview & Syllabus
- Databases and Database Applications
- Layers
- The Relational Model
- SQL - The Basics



Course Objectives

- Apply the entity-relationship and relational data models, as well as data structures, integrity constraints, and operations for each data model, using the relational language SQL
- Describe the theory of database design, including functional dependencies, normal forms, and dependency preservation.
- Develop SQL integration with Java/Python programs.
- Understand how database engines and storage systems impact database system performance
- Perform query processing and optimization.

Course Information

Live weekly class session
Wednesdays 6:30-8:30 PM EDT
Office Hours weekly session
Fridays 8:00-9:00 PM EDT
Zoom Link - Passcode database
Sign-up link for Office Hours

Please sign up for office hours if you will be attending - you can join any time regardless of the time selected

Week	Topic	Project	Quiz
1	Database and querying basics	Project 1: Create and query	
2	Normalization and joins	Project 2: From one table to many	Quiz 1
3	Stored procedures, and aggregation	Project 3: Functions and stored procedures	
4	Connecting databases to applications	Project 4: Building a CLI application	Quiz 2
5	View Layer Deep Dive	Project 5: GUI application	
6	Users, Authorization, and Storage	Project 6: Final project plan	Quiz 3
7	B+ Trees and Indexing	Final Project: Database	
8	Query Plans and Optimization	Final Project: Application	Final Quiz

Grading & Academic Integrity

Activity	Percentage
Projects (6)	30%
Quizzes (3)	20%
Final Quiz	20%
Final Database	15%
Final Application	15%

Acceptable	Unacceptable
Attending tutoring or office hours to improve your assignments	Submitting work completed by someone else
Using AI to explain course concepts or errors or act as tutor	Using AI to write all or a portion of your assignments
Using AI to create test data	Using AI to create images or ER diagrams
Adapting small pieces of code from other sources with proper attribution	Copying code directly from a source without attribution or copying large amounts of code

Main Text

Textbook:
Database System Concepts
7th edition
Abraham Silberschatz, Henry F. Korth, S. Sudarshan
ISBN: 978-0078022159

Supplements

- SAMS Teach Yourself SQL in 10 Minutes
  - For absolute beginners
- SAMS Teach Yourself SQL in 24 Hours
  - More technical
- MySQL Crash Course (Silva)
  - Resource specific to MySQL

Other Important Information

- Projects due Saturdays, Quizzes due Tuesdays, Final Quiz due last Friday
- I will grade on Sundays so you have immediate feedback (exception: Nov 9)
- Certain assignments require a starter .sql file. These will unlock when you hand in the previous assignment. Make sure you download and use the starter file as a base. If you copy/paste, there may be errors.
- Late penalty of 3 points per day applied automatically by Canvas. You start with 6 points as a buffer.
- If you cannot get MySQL workbench or alternative working by the end of Week 1, you should strongly consider dropping during add/drop. Students who struggled with this have a difficult time finishing the course.
- If you are are more than two week behind at the deadline to withdraw, I recommend withdrawing. Incompletes will only be granted for the Final Project.

Questions?

What is a Database?

A collection of Data.

How do we store this data?

How do we access this data?

How do we define this data?

How do we interact with this data?

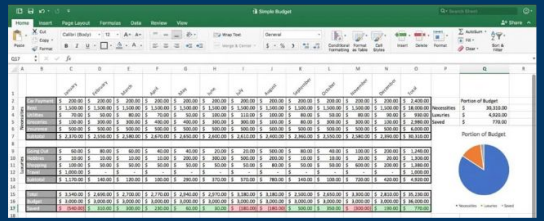
How do we manipulate this data?

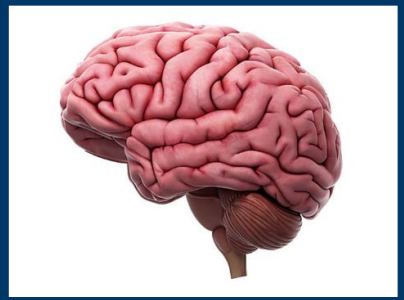
How do we trust this data?

How does this data relate to itself?

What is a Database?

How do we store this data?





What is a Database Management System (DBMS)?

A collection of interrelated data and a set of programs to access those data.

Primary Goal:
Store and retrieve database information that is both convenient and efficient.

How do we access this data?

What is a Database Management System (DBMS)?

How do we define this data?

Relational Database Management System (RDBMS)

Object-Relational Database Management System (ORDBMS)

NoSQL Database Management System (NoSQL)

What is PostgreSQL?
Document Databases

From Database to Application

How do we interact with this data?

IMDb

From Database to Application

How do we manipulate this data?

Create

Read

Update

Delete

COURSE SEARCH - Course Search

Academic Year 2024 - 2025 - Spring

Course Search

Course Code

Begins With | csc6

Course Title

Begins With

Division

Instructor

Search...

Department

Search...

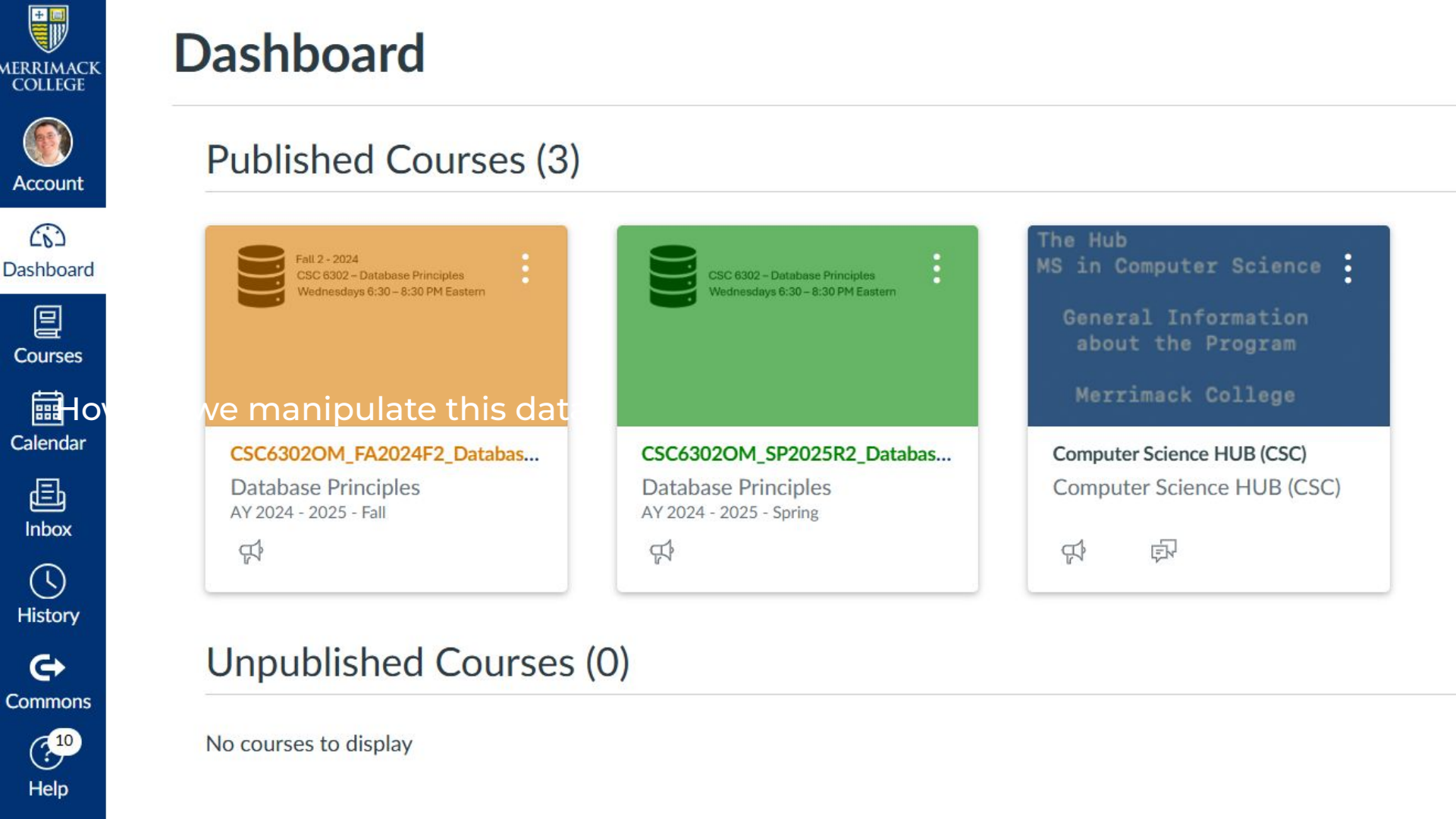
Meeting Type

Meets on Selected Days

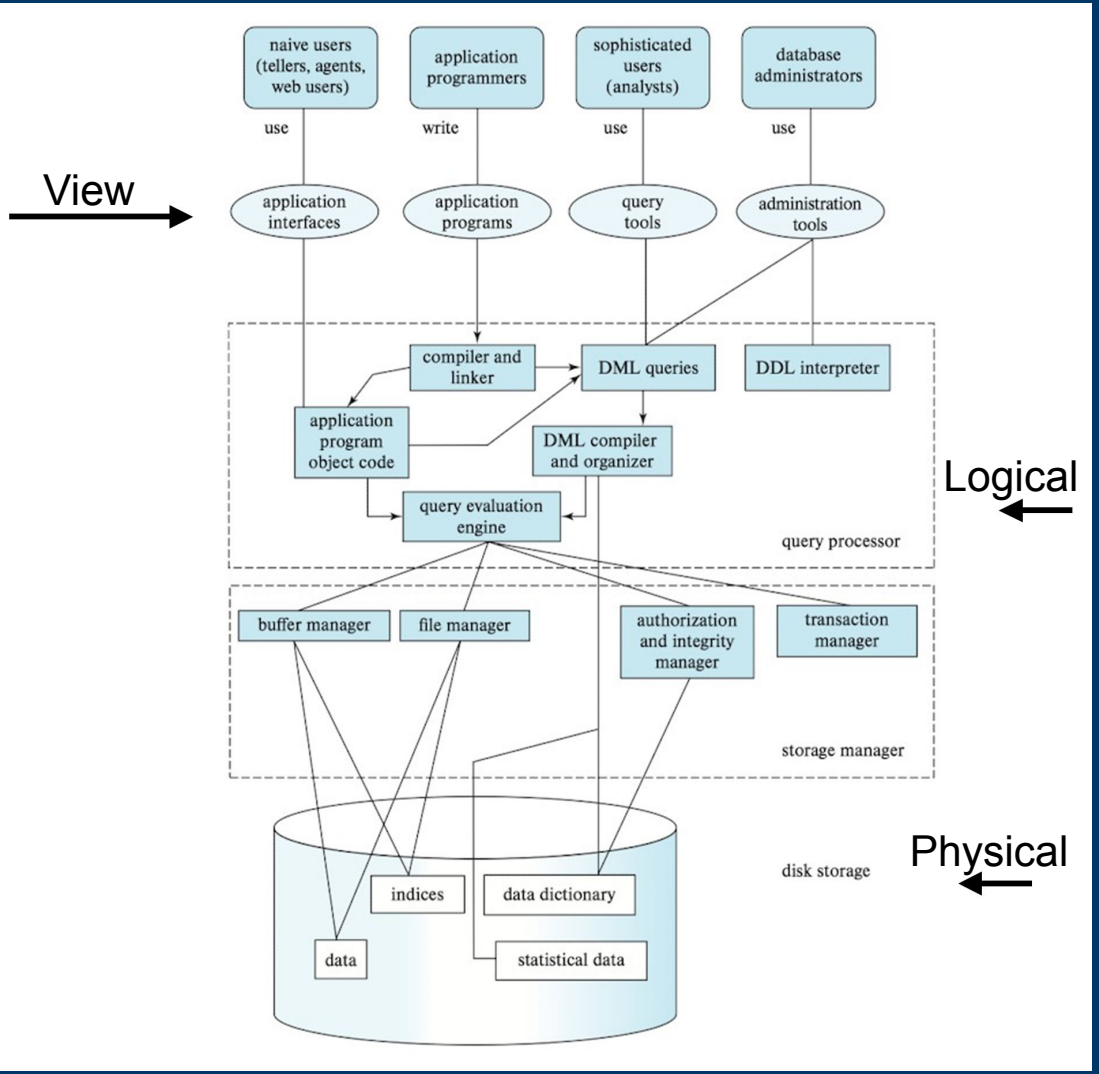
M Tu W Th F Sa Su

Search Courses

Course Code	Title	Faculty	Seats Open	Status	Schedule	Credits
CSC-6000-OM	Basic Programming and Discrete M	Paulo Fernandes Paulo Fernandes Paulo Fernandes	65/99	Open	**Thu** 6:30-8:30 PM 1/15/2025 - 3/7/2025 Online Online - OL	4.00
CSC-6000-ON	Basic Programming and Discrete M	Paulo Fernandes Paulo Fernandes Paulo Fernandes	80/99	Open	**Thu** 6:30-8:30 PM 3/17/2025 - 5/9/2025 Online Online - OL	4.00
CSC-6003-OM	Foundations of Programming	Ares Manuel Ca... Ares Manuel Ca... Ares Manuel Ca... Ares Manuel Ca... Ares Manuel Ca...	79/99	Open	**Thu** 6:30-8:30 PM 1/15/2025 - 3/7/2025 Online Online - OL	4.00
CSC-6003-OO	Foundations of Programming	Ares Manuel Ca... Ares Manuel Ca... Ares Manuel Ca... Ares Manuel Ca... Ares Manuel Ca... Sergei Pustylnikov Sergei Pustylnikov Sergei Pustylnikov	68/99	Open	**Thu** 6:30-8:30 PM 3/17/2025 - 5/9/2025 Online Online - OL	4.00
CSC-6013-OM	Algorithms + Discrete Structures	Paulo Fernandes Paulo Fernandes Paulo Fernandes	77/99	Open	**Mon** 6:30-8:30 PM 1/15/2025 - 3/7/2025 Online	4.00


MERRIMACK COLLEGE
Account
Dashboard
Courses
Calendar
Inbox
History
Commons
10
Help
Dashboard
Published Courses (3)
Fall 2 - 2024
CSC 6302 – Database Principles
Wednesdays 6:30 – 8:30 PM Eastern
CSC6302OM_FA2024F2_Databas...
Database Principles
AY 2024 - 2025 - Fall
CSC 6302 – Database Principles
Wednesdays 6:30 – 8:30 PM Eastern
CSC6302OM_SP2025R2_Databas...
Database Principles
AY 2024 - 2025 - Spring
The Hub
MS in Computer Science
General Information about the Program
Merrimack College
Computer Science HUB (CSC)
Computer Science HUB (CSC)
Unpublished Courses (0)
No courses to display

Application Layers



ACID

How do we trust this data?

Atomicity:
All or nothing

Consistency:
Valid state transition

Isolation:
Concurrent transactions do not interfere

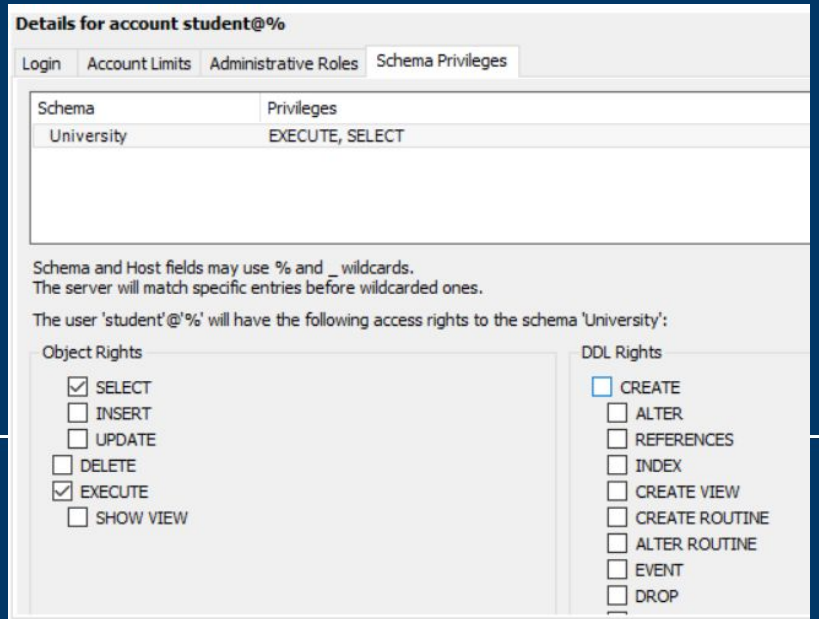
Durability:
Permanent changes once committed

Authorization & Authentication

How do we trust this data?

Authentication:
Who are you?

Authorization:
What can you do?



Relational Model

How does this data relate to itself?

Common Term	Relational Term
Table	Relation
Column	Attribute
Type	Domain
Row	Tuple

Some Data Types

Numeric:
- INT, FLOAT, DECIMAL

String:
- CHAR, VARCHAR, TEXT

Date/Time:
- DATE, DATETIME, TIMESTAMP, TIME

Other:
- ENUM, BLOB

https://dev.mysql.com/doc/refman/8.4/en/data-types.html

Special Considerations

Boolean values:

TINYINT -> 1 and 0 values for True or False

NULL:

Indicates no value is present
Different from “” (empty string)
Remember this when importing data

https://dev.mysql.com/doc/refman/8.4/en/data-types.html

COURSE SEARCH - Course Search

Academic Year 2024 - 2025 - Spring

Course Search

Course Code

Begins With | csc6

Course Title

Begins With

Division

Instructor

Search...

Department

Search...

Meeting Type

Meets on Selected Days

M Tu W Th F Sa Su

Search Courses

Course Code	Title	Faculty	Seats Open	Status	Schedule	Credits
CSC-6000-OM	Basic Programming and Discrete M	Paulo Fernandes Paulo Fernandes Paulo Fernandes	65/99	Open	**Thu** 6:30-8:30 PM 1/15/2025 - 3/7/2025 Online Online - OL	4.00
CSC-6000-ON	Basic Programming and Discrete M	Paulo Fernandes Paulo Fernandes Paulo Fernandes	80/99	Open	**Thu** 6:30-8:30 PM 3/17/2025 - 5/9/2025 Online Online - OL	4.00
CSC-6003-OM	Foundations of Programming	Ares Manuel Ca... Ares Manuel Ca... Ares Manuel Ca... Ares Manuel Ca... Ares Manuel Ca...	79/99	Open	**Thu** 6:30-8:30 PM 1/15/2025 - 3/7/2025 Online Online - OL	4.00
CSC-6003-OO	Foundations of Programming	Ares Manuel Ca... Ares Manuel Ca... Ares Manuel Ca... Ares Manuel Ca... Ares Manuel Ca... Sergei Pustylnikov Sergei Pustylnikov Sergei Pustylnikov	68/99	Open	**Thu** 6:30-8:30 PM 3/17/2025 - 5/9/2025 Online Online - OL	4.00
CSC-6013-OM	Algorithms + Discrete Structures	Paulo Fernandes Paulo Fernandes Paulo Fernandes	77/99	Open	**Mon** 6:30-8:30 PM 1/15/2025 - 3/7/2025 Online	4.00

COURSE SEARCH - Course Search

ENUM('Open', 'Closed')

CHAR(11) VARCHAR(50) DECIMAL(4,2)

Academic Year 2024 - 2025 - Spring

Course Search

Course Code

Begins With | csc6

Course Title

Begins With

Division

Instructor

Search...

Department

Search...

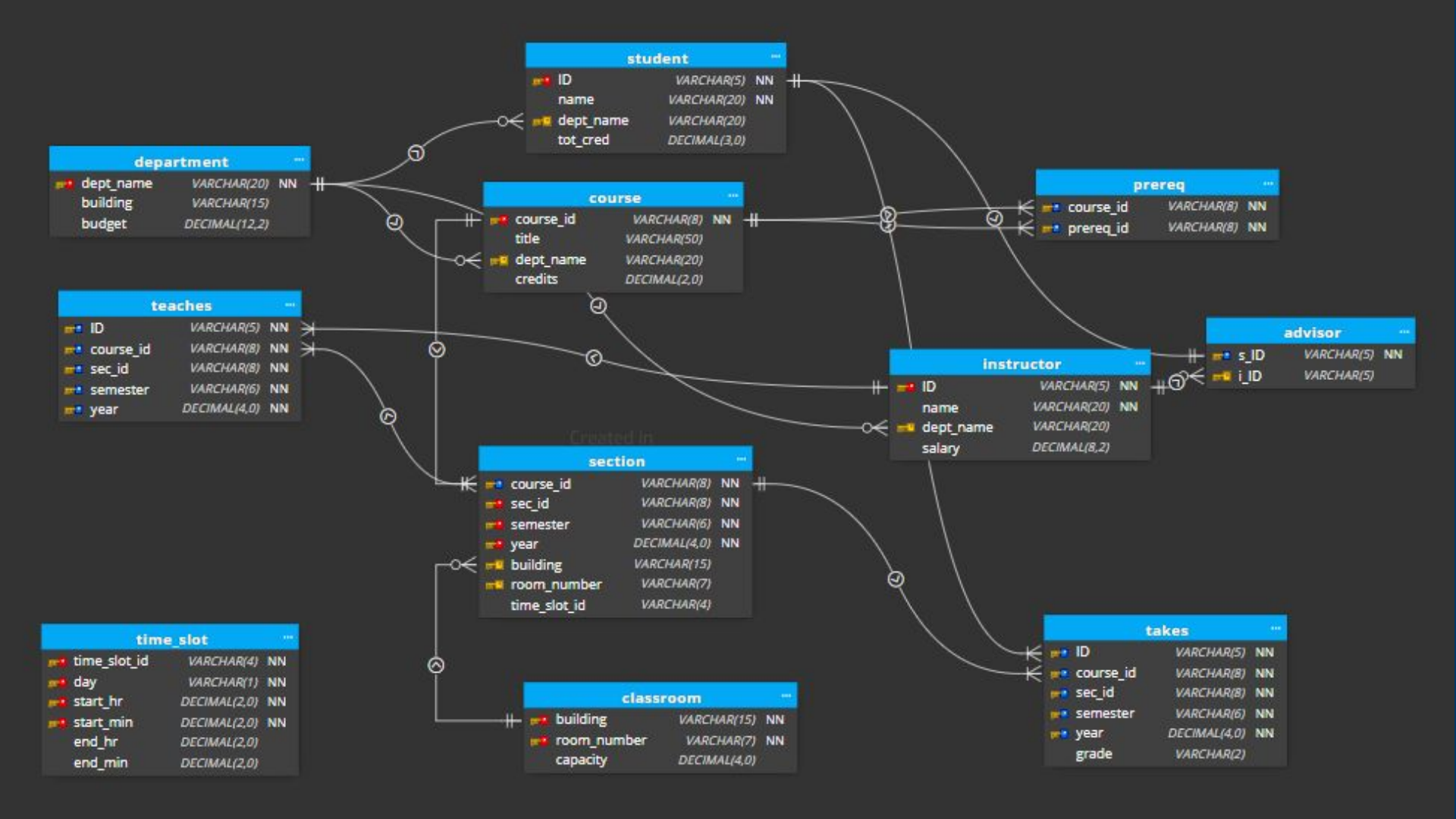
Meeting Type

Meets on Selected Days

M Tu W Th F Sa Su

Search Courses

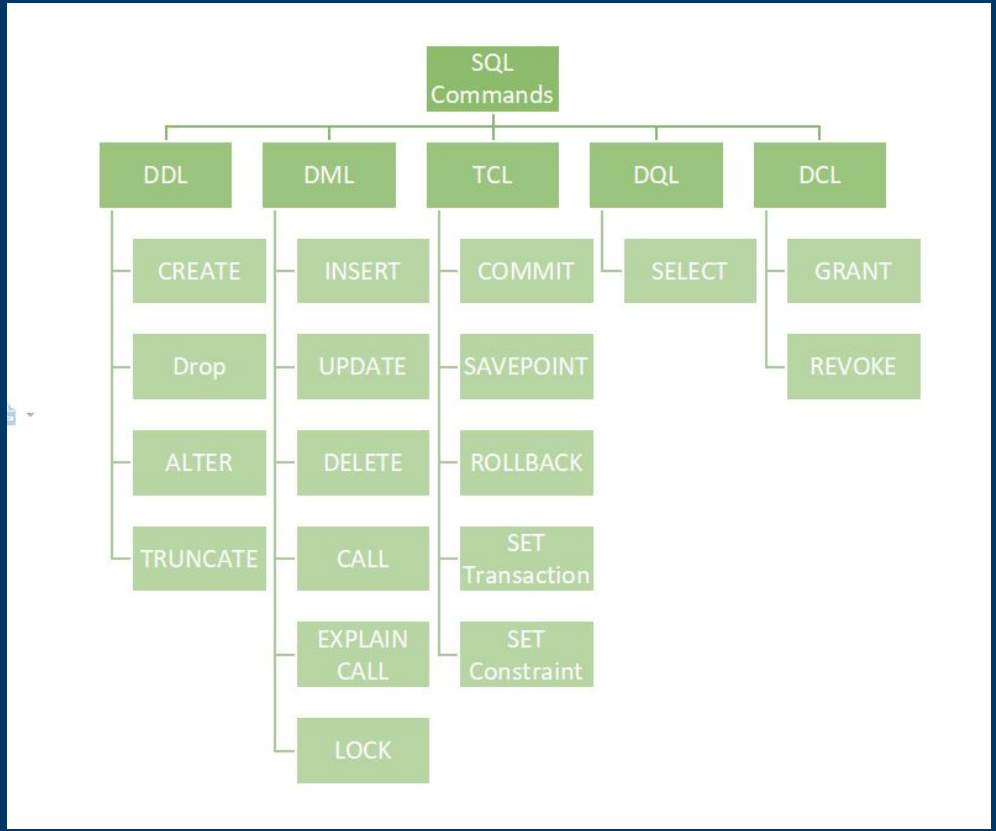
Course Code	Title	Faculty	Seats Open	Status	Schedule	Credits
CSC-6000-OM	Basic Programming and Discrete M	Paulo Fernandes Paulo Fernandes Paulo Fernandes	65/99	Open	**Thu** 6:30-8:30 PM 1/15/2025 - 3/7/2025 Online Online - OL	4.00
CSC-6000-ON	Basic Programming and Discrete M	Paulo Fernandes Paulo Fernandes Paulo Fernandes	80/99	Open	**Thu** 6:30-8:30 PM 3/17/2025 - 5/9/2025 Online Online - OL	4.00
CSC-6003-OM	Foundations of Programming	Ares Manuel Ca... Ares Manuel Ca... Ares Manuel Ca... Ares Manuel Ca... Ares Manuel Ca...	79/99	Open	**Thu** 6:30-8:30 PM 1/15/2025 - 3/7/2025 Online Online - OL	4.00
CSC-6003-OO	Foundations of Programming	Ares Manuel Ca... Ares Manuel Ca... Ares Manuel Ca... Ares Manuel Ca... Ares Manuel Ca... Sergei Pustylnikov Sergei Pustylnikov Sergei Pustylnikov	68/99	Open	**Thu** 6:30-8:30 PM 3/17/2025 - 5/9/2025 Online Online - OL	4.00
CSC-6013-OM	Algorithms + Discrete Structures	Paulo Fernandes Paulo Fernandes Paulo Fernandes	77/99	Open	**Mon** 6:30-8:30 PM 1/15/2025 - 3/7/2025 Online	4.00


student
ID VARCHAR(5) NN
name VARCHAR(20) NN
dept_name VARCHAR(20)
tot_cred DECIMAL(3,0)
department
dept_name VARCHAR(20) NN
building VARCHAR(15)
budget DECIMAL(12,2)
course
course_id VARCHAR(8) NN
title VARCHAR(50)
dept_name VARCHAR(20)
credits DECIMAL(2,0)
prereq
course_id VARCHAR(8) NN
prereq_id VARCHAR(8) NN
teaches
ID VARCHAR(5) NN
course_id VARCHAR(8) NN
sec_id VARCHAR(8) NN
semester VARCHAR(6) NN
year DECIMAL(4,0) NN
instructor
ID VARCHAR(5) NN
name VARCHAR(20) NN
dept_name VARCHAR(20)
salary DECIMAL(8,2)
advisor
s_ID VARCHAR(5) NN
i_ID VARCHAR(5)
section
course_id VARCHAR(8) NN
sec_id VARCHAR(8) NN
semester VARCHAR(6) NN
year DECIMAL(4,0) NN
building VARCHAR(15)
room_number VARCHAR(7)
time_slot_id VARCHAR(4)
time_slot
time_slot_id VARCHAR(4) NN
day VARCHAR(1) NN
start_hr DECIMAL(2,0) NN
start_min DECIMAL(2,0) NN
end_hr DECIMAL(2,0)
end_min DECIMAL(2,0)
classroom
building VARCHAR(15) NN
room_number VARCHAR(7) NN
capacity DECIMAL(4,0)
takes
ID VARCHAR(5) NN
course_id VARCHAR(8) NN
sec_id VARCHAR(8) NN
semester VARCHAR(6) NN
year DECIMAL(4,0) NN
grade VARCHAR(2)

SQL Basics

Structured Query Language
- Syntax varies slightly between platforms (MySQL vs T-SQL, etc)

DDL – Data Definition Language
DQL – Data Query Language
DML – Data Manipulation Language
DCL – Data Control Language
TCL – Transaction Control Language



https://www.geeksforgeeks.org/sql-ddl-dql-dml-dcl-tcl-commands/

Database Creation

```
DROP DATABASE University;

CREATE DATABASE IF NOT EXISTS
    University;

USE University;
```

For the purposes of this course, avoid errors by dropping the database at the top of the file and adding ‘if [not] exists’ language

Use the USE command to select your schema.

Commands are not case sensitive, but it is conventional to ALLCAPS commands to facilitate reading the sql

Table Creation

```
CREATE TABLE IF NOT EXISTS course
      (course_id
            VARCHAR(8),
       title
            VARCHAR(50),
       dept_name
            VARCHAR(20),
       credits
            NUMERIC(2,0) CHECK (credits
            > 0));
```

```
create table if not exists instructor
      (ID
            varchar(5),
       name
            varchar(20) not null,
       dept_name
            varchar(20),
       salary
            numeric(8,2) check (salary >
            29000));
```


Terminate statements with a semicolon;

Table Insertion

```
create table if not exists course
        (course_id
                varchar(8),
         title
                varchar(50),
         dept_name
                varchar(20),
         credits
                numeric(2,0) check
                (credits > 0));
```

```
INSERT INTO
course (course_id, title, dept_name, credits)
VALUES
('CSC6302', 'Database Principles', 'Comp. Sci.', 4.0),
('CSC6013', 'Discrete Structures', 'Comp. Sci.', 4.0);

-- Double dash for a single line comment
/* Multi-line comments
*/
```

Table Insertion from CSV

You may need to run the statement SHOW VARIABLES LIKE "secure_file_priv"; to find the directory that MySQL will allow uploads from. Place your csv there.

The file path should always use forward slashes, even on Windows.

```
LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL
Server 8.3/Uploads/courses.csv'
INTO TABLE University.courses FIELDS
TERMINATED BY ','
ENCLOSED BY "" LINES TERMINATED BY '\n'
IGNORE 1 Rows;
-- Ignore rows if you have header columns
```

First Queries

```
create table instructor
        (ID
                varchar(5),
         name
                varchar(20) not null,
         dept_name
                varchar(20),
         salary
                numeric(8,2) check (salary >
                29000));
```

SELECT:
The attributes or values to ‘return’
FROM:
Selects the relation to get the data from
WHERE:
Criteria used to ‘filter’ tuples

```
SELECT ID
FROM instructor
WHERE dept_name = ‘Comp. Sci.’
```

String Matching

Patterns are case sensitive when using =.
LIKE is a similar match (not case sensitive).
‘_’ is a single character wildcard.
‘%’ is a 0 or more character wildcard.

Pattern matching examples:

'Intro%' matches any string beginning with ‘Intro’.
'%Comp%' matches any string containing ‘Comp’ as a substring.
'_ _ _' matches any string of exactly three characters.
'_ _ _ %' matches any string of at least three characters.

Questions?

Project 1:

Congratulations! You have been hired to create a database application for Merrimack River Cruises.

This local business is currently using a spreadsheet to keep track of the voyages of their small but enthusiastic customer base. Your first task is to take that .csv and turn it into a database table.