**DASC 5333**

**1/14/2025**

Self Annotation

**DASC 5333 Database Systems for Data Science  
CSCI 4333 Design of Database Systems  
Spring 2025  
General Information and Course Policies**

**1. General Information**

22183 CSCI 4333.1 Design Of Database Systems MW 1:00-2:20 Delta 241  
22178 DASC 5333.1 Database Systems for Data Science T 1:00-3:50 Delta 237

This is a first course in database for both BS CS students and MS DS student.

**1.1 Instructor**

Dr. Kwok-Bun Yue, Professor of Computer Science and Computer Information Systems, Chair, Data Science  
Delta 163, 281-283-3864, yue at uhcl.edu; URL: <http://dcm.uhcl.edu/yue/>

My regular office hour will be held on 1/14/2025 to 4/28/2025: MW 2:25PM to 3:45PM and Tuesday 3:50-4:10PM. Office hours will be conducted in person (Delta 163) and upon request, via [Zoom meeting: 616 099 762](https://uhcl.zoom.us/j/616099762?pwd=TlhGOXNrQjlCTVhkSTVEeTdBUDBpUT09). You can schedule a meeting with me outside my office hours by sending an email to me: yue @ uhcl dot edu. You are encouraged to communicate your questions with me through email. I usually respond quick.

**1.2 Teaching Assistant**

**1.2.1 CSCI 4333.1, CSCI 4333.2 and DASC 5333**

Pavan Kumar Kodavali

A person with a beard and mustache

Description automatically generated

For regular correspondence with the TA, send it to UHCL Email Id: [kodavali@uhcl.edu](mailto:kodavali@uhcl.edu). Set up the UHCL spam filter server for your UHCL account to accept this email address as an approved sender. Otherwise, your email may be quarantined by the spam filter server. If you want me to be aware of any particular communications with the TA, you may copy the email to me.

Tentative TA Office hours:

|  |  |
| --- | --- |
| **Day** | **TA Office Hours** |
| Monday | 10:00 AM - 1:00 PM |
| Tuesday | 10:00 AM - 1:00 PM |
| Wednesday | 11:00 AM - 1:00 PM |
| Thursday | 10:00 AM - 1:00 PM & 4:00 PM to 7:00 PM |

Pavan will be stationed in the Delta Lab during his office hours. You may also request Zoom for TA: <https://uhcl.zoom.us/my/pavankodavali>

**1.3 Laboratory Administrations**

You may address account and software problems of the DCM server to the systems administrator, [Ms. Krishani Abeysekera](mailto:abeysekera@uhcl.edu). and her assistants. Always copy your email to me.

**1.4 Other Useful Information**

* UHCL emergency hot line (to check weather related closing, for example): 281-283-2221.

**1.5 Textbooks (Recommended, Optional)**

Ricardo., Katherine, & Urban, Susan (2015)*Databases Illuminated*, 3rd Edition, Jones & Bartlett, Mississauga, Ontario, Canada.

**1.6 Course Description**

**CSCI 4333:**

From Catalog: Prerequisite: CSCI 2315. Design of database systems, data description and manipulation languages, data models, entity-relationship model, relational model, SQL, relational algebra, normalization theory, DBMS, Internet, data base design, data flow diagrams and implementation of data base systems. Laboratory instruction.

**DASC 5333:**

From Catalog: Design of database systems, data definition and manipulation languages, data models, entity-relationship model, relational model, SQL, relational algebra, normalization theory, DBMS, Internet, database implementation. Focus on applying DB theory and practice to support data science applications. Laboratory instruction. Prerequisites: DASC 5032 or equivalence

**1.7 Student Learning Outcomes (SLO)**

**CSCI 4333:**

After completing the course, the students are expected to be able to

1. Describe the stages of database design, various database architectures and data models.
2. Explain the concepts of the entity-relationship model, and the relational model.
3. Explain the theoretical background of relational database, including relational algebra and relational calculus.
4. Implement relational database systems using DBMS, SQL, embedded SQL, including both data definition and manipulation languages
5. Explain the importance of normalization of databases, and convert a given relational database into different normal forms.

**DASC 5333:**

After completing the course, the students are expected to be able to

1. Describe the fundamental concepts of database systems.
2. Apply simple data modeling techniques using the UML class diagram and the relational model.
3. Explain the features and theory of the relational model.
4. Implement relational database systems for data science applications using relational DBMS, SQL, and embedded SQL with Python.
5. Explain the importance of normalization of relational databases, and convert a given relational database to appropriate normal forms.
6. Explain the concepts of NoSQL databases and implement simple NoSQL database solutions.

**1.8 Prerequisites**

The following courses or their equivalent are required:

* CSCI 4333: CSCI 2315 Data Structures
* DASC 5333: DASC 4301 Python Programming for Data Science, or equivalent

Languages: The course uses SQL, Python and (MongoDB Query Language (MQL) with Javascript, and/or Cypher (for Neo4j)). No prior SQL, MQL or Cypher language knowledge is assumed. Students are expected to know an object-oriented language, such as Python, Java, C# or C++. Proficiency in Python is important in data science in general, and this course in particular.

**1.9 Course Format**

Traditional lectures, homework and programming assignments.

**2. Course Policies and Guidelines**

Please see: <http://dcm.uhcl.edu/yue/course_policy.html>

**3. Grading Policy**

Grades will be assigned based *solely* on homework and examination scores. *No other factors will be considered.* In particular, students have requested me to reconsider their grades using the following reasons in the past:

1. Expected a higher grade
2. Good course participation
3. Good improvement during the semester; better final grades than mid-term grades
4. Have put in extra efforts
5. Need to avoid probation
6. Financial needs
7. Loss of scholarship
8. Loss of job opportunity
9. Loss of practical training opportunity
10. Avoid probation; avoid suspension
11. Need to graduate
12. Company relocation
13. Immigration status needs
14. Family needs
15. Sickness during the semester
16. and many others.

These requests had all been declined politely but firmly in the past.

There will also be *no* 'special project' that you can work on to improve your grades after the final examination. Anything I offer to one student will be offered to the entire class.

The total score is computed using the following percentages:

Homework: 30%  
Mid-term Exam: 30%  
Final Exam: 40%

Last Day to Drop/Withdraw: *April 7, 2025 (Monday)*

**Grade Assignment Table**

|  |  |
| --- | --- |
| [92..100] | A |
| [90..92) | A- |
| [87..90) | B+ |
| [83..87) | B |
| [80..83) | B- |
| [77..80) | C+ |
| [73..77) | C |
| [70..73) | C- |
| [67..70) | D+ |
| [63..67) | D |
| [60..63) | D- |
| [0..60) | F |

*ask a lot of questions*

a penalty of *25% deduction* per *week day*

HW #1 sol:

<https://dcm.uhcl.edu/yue/courses/joinDB/Fall2024/hw/h1sol.sql.txt>

E.g.

-- (a) List the stuId, names, major departments and minor code of all students minoring in CSCI, CINF or ITEC in the following manner.

SELECT DISTINCT s.stuId, CONCAT(s.fname, ' ', s.lname) AS student,

d.deptName AS major,

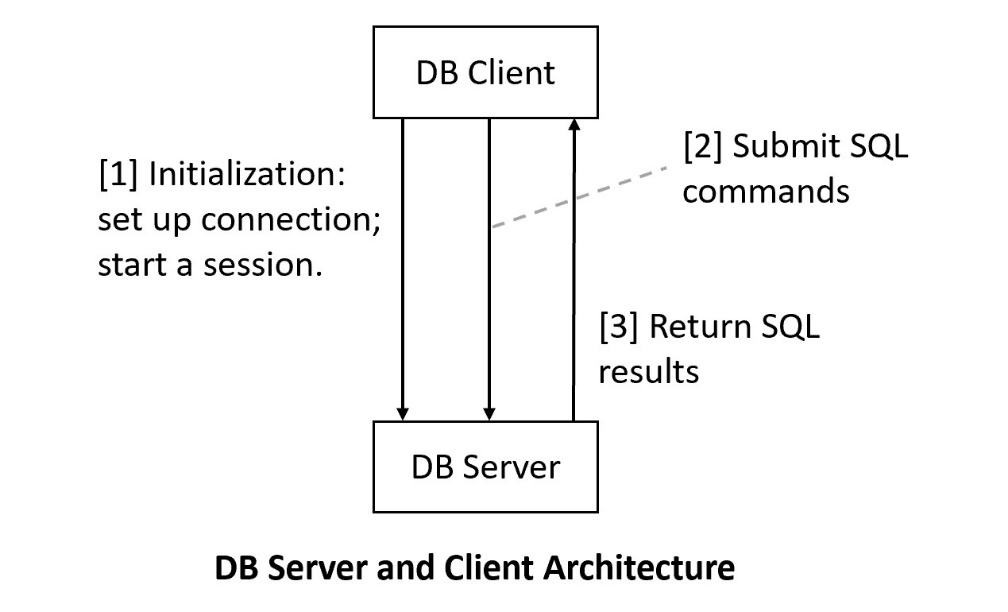
s.minor

FROM student AS s INNER JOIN department AS d ON (s.major = d.deptCode)

WHERE s.minor = 'CSCI'

OR s.minor = 'CINF'

OR s.minor = 'ITEC';



XAMPP provides the MYSQL Server.

Install XAMPP: development package for server software.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Green: MYSQL server is running and listening at port 3306.

MySQL Prompt (MySQL client): command line client.

Working directory: e.g. C:\S2025\_JointDB\website\demo\d5333

Terminal in the working directory

Text editor: notepad++

MySQL client:

mysql -u yue -p

**1.1 Installing Toyu DB in your computer**

**Installing MySQL:**

1. XAMPP: contains MariaDB (close to MySQL): <https://www.apachefriends.org/>
   1. Include other server software, e.g., Apache (the Web server needed in a Python-Web application assignment)
   2. Include MySQL command line client (mysql console/prompt)
   3. Does not include MySQL Workbench (a MySQL client)
2. MySQL: <https://www.mysql.com/>
   1. Include MySQL command line client
   2. Include MySQL Workbench

A computer screen with white text

Description automatically generated

In MYSQL: schema = database

A screen shot of a computer

Description automatically generated

Set path to include the client software:

C:\xampp\mysql\bin

Allosws you to run: mysql -u yue -p

Instead of

C:\xampp\mysql\bin\mysql -u yue -p

A screenshot of a computer

Description automatically generated

GIU client: HeidiSQL

A screenshot of a computer

Description automatically generated

[1] List the stuId, names, major departments and minor code of all students

minoring in CSCI, CINF or ITEC in the following manner.

+--------+-----------------+------------------------------+-------+

| stuId | student | major | minor |

+--------+-----------------+------------------------------+-------+

| 100000 | Tony Hawk | Computer Science | CINF |

| 100001 | Mary Hawk | Computer Science | CINF |

| 100002 | David Hawk | Computer Science | ITEC |

| 100003 | Catherine Lim | Information Technology | CINF |

| 100006 | Lillian Johnson | Computer Information Systems | ITEC |

| 100009 | Linda King | Arts | CSCI |

+--------+-----------------+------------------------------+-------+

6 rows in set)

[1] (b) List the stuId, fnames, lname, major code and minor code of all students

Analysis:

[1] Output columns:

1. stuId
2. fname
3. lname
4. major
5. minor

[2] Sources: student

[3] Conditions: none; all students

SELECT DISTINCT – [1]   
FROM – [2]  
WHERE -- [3]

[1] (c) List the stuId, fnames, lname, major code and minor code of all students

Analysis:

[1] Output columns: label: value

1. stuId
2. student: fname concatenate with ‘space’ and lname, CONCAT(fname, ' ', lname), e.g: Tony Hawk:
3. major
4. minor

[2] Sources: student

[3] Conditions: none; all students

[1] (c) List the stuId, fnames, lname, major code and minor code of all students

Analysis:

[1] Output columns: label: value

1. stuId
2. student: fname concatenate with ‘space’ and lname, CONCAT(fname, ' ', lname), e.g: Tony Hawk:
3. major: department.deptName
4. minor

[2] Sources:

1. student
2. department

[3] Conditions:

* Join condition: student.major = department.deptCode