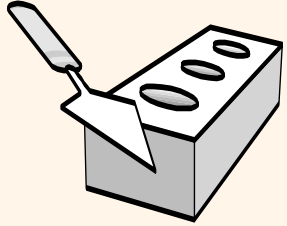


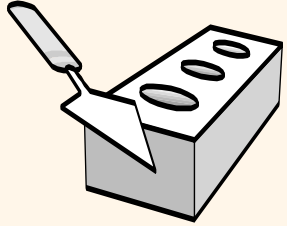
COMP7640
Database Systems and Administration

Instructor: Dr. Renchi YANG

Teaching Assistants:

ZHAO Yunxiang (csyxzhao@comp.hkbu.edu.hk)

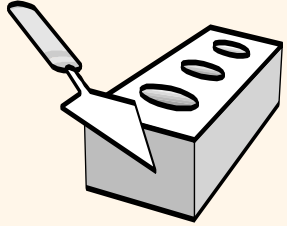
YI Lanjing (csljyi@comp.hkbu.edu.hk)



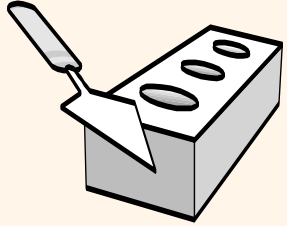
About the Course Instructor

- Assistant Professor in the Dept. of CS, HKBU
  - Jan 2023 - Present
- Educational Background:
  - Ph.D., Nanyang Technological Univ., Singapore
  - B.Eng., Beijing Univ. of Posts & Telecommunications
- Research Interests
  - Big Data Management, Analytics & Mining
- Homepage: https://www.comp.hkbu.edu.hk/~renchi/
- Email: renchi@hkbu.edu.hk
- Office: DLB 644

In-class Rules

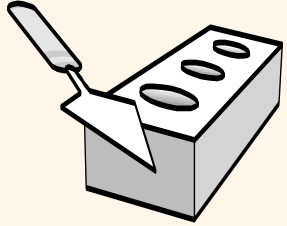


- Turn off your cell phone or keep it on vibrate/silent mode
- Keep listening, thinking, taking notes, …
  - Interrupt me anytime if I'm speaking too fast/you don't understand the concept
- Try to work on the in-class exercises (but not simply wait for the answers).
- But **NO TALKING PLEASE!**
  - To respect your fellow students, instructor & yourself.
- Feel free to come to me in class breaks
  - Better to speak English as other students might not understand Mandarin or Cantonese



Out-of-class Q&A

- Email correspondence
  - Send your questions to TAs first as they are more responsive. Your questions will be forwarded to me if they can not resolve them.
    - TAs: Mr. ZHAO Yunxiang (csyxzhao@comp.hkbu.edu.hk) and Mr. YI Lanjing (csljyi@comp.hkbu.edu.hk)
- In-person meetings/discussions
  - Office (DLB 644) Hours: 2:00 – 5:30 PM Tuesday
  - Make an appointment with me first
  - Better to come in group

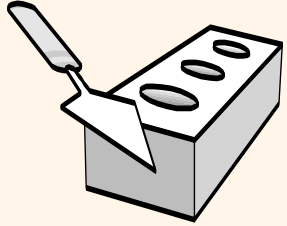


Tentative Schedule:

1st Week (Jan 16): 8:00-9:20 AM @ WLB 103

Other Weeks: Thu 6:30 PM-9:20 PM @ WLB 103

Lecture	Date	Topic	Assignment	Project
Lecture 1	Jan 16	**Introduction & ER Model**	**Assignment #1 Due: Feb 20**	**Group Project Release: Feb 20 Due: Apr 10**
Lecture 2	Jan 23	**ER & Relational Model**		
Lecture 3	Feb 6	**Relational Model**		
Lecture 4	Feb 13	**SQL**		
Lecture 5	Feb 20	**SQL**		
Lecture 6	Feb 27	**Functional Dependencies & Normalization**		



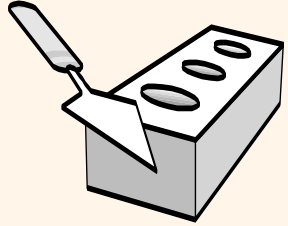
Tentative Schedule:

1st Week (Jan 16): 8:00-9:20 AM @ WLB 103

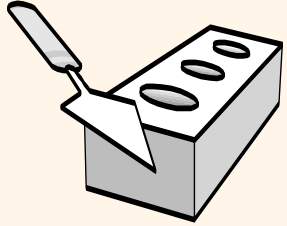
Other Weeks: Thu 6:30 PM-9:20 PM @ WLB 103

Lecture	Date	Topic	Assignment
Lecture 7	Mar 6	**Storage Management & Access Methods**	**Assignment #2** **Due: Mar 27**
Lecture 8	Mar 13	**Tree-based Index**	
Lecture 9	Mar 20	**Hash Index**	**Assignment #3** **Due: Apr 24**
Lecture 10	Mar 27	**Query Evaluation**	
Lecture 11	Apr 3	**Query Optimization**	
Lecture 12	Apr 10	**Transaction Management & Concurrency Control**	
Lecture 13	Apr 24	**Crash Recovery**	

Assessment



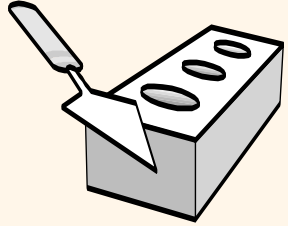
- Continuous assessment: 40%
  - 3 written assignments: 24% (8% for each assignment)
  - Group project: 16%
    - Form your group as early as possible
- Final exam: 60%



Group Project

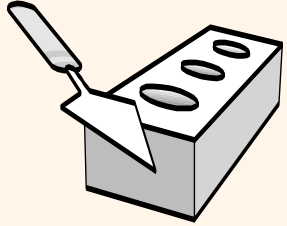
- Group Size:
  - 4~6 members
  - Each group should have a group leader
- Group Enrollment
  - Start: 12:00 AM Feb 6
  - Due: 11:59 PM Feb 27
- Each member of the same group will receive the same marks.

Assessment Guides



- Criterion-referenced assessment
  - Intended Learning Outcomes & Rubrics
- To pass this subject:
  - Final exam score must be ≥ 30%
  - Overall score must be ≥ 35%
- No LATE submission is accepted.

BEWARE!!!

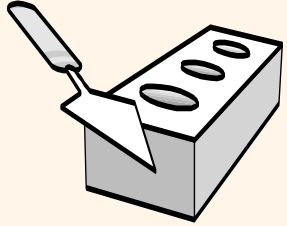


- Unless otherwise stated, all work submitted by you should be your own.

- **Copying or sharing of assignments or any submitted work constitutes cheating**.

- If there is any doubt about the appropriateness of your actions, please contact the instructor for explicit clarification.

Cost of Plagiarism

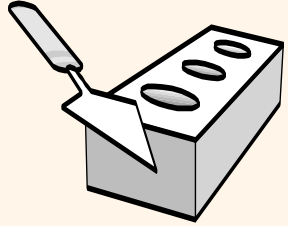


- Plagiarism is an offense and will result in appropriate disciplinary action against those involved.

- Penalty will be applied indiscriminatingly among those who involve (the one who copy and the one being copied). The *minimum penalty* would be receiving zero mark for the submitted work.

- Please refer to the following URL for the university's guideline on penalizing plagiarism:

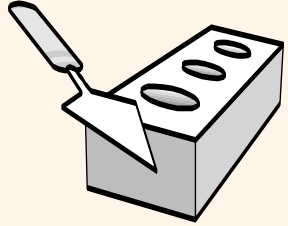
https://ar.hkbu.edu.hk/quality-assurance/university-policy-and-guidelines/academic-integrity/section-2-plagiarism



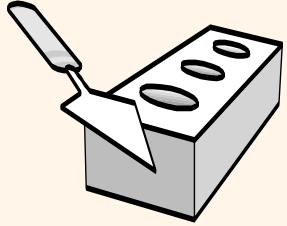
Usage of Generative AI

- Proper Use of Generative AI
  - explaining or clarifying concepts;
  - demonstrating and guiding practices of techniques;
  - planning and brainstorming on projects;
  - giving feedback on drafts;
  - generating samples for discussion and critical review.

University's guideline: https://bba.hkbu.edu.hk/academics/teaching-and-learning-supports



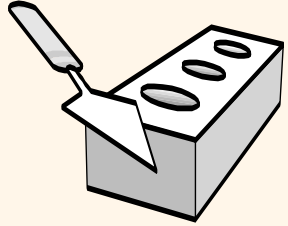
Overview



What is a Database?

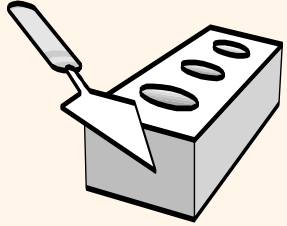
- A database is a collection of data that is organized so that its contents can be easily accessed, managed, and updated.

Why do we need a Database?



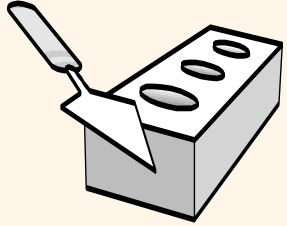
- Submit a course add/drop form to AR.
- Pay tuition fees in Hang Seng Bank.
- Borrow the textbook from the library.
- Order the textbook from e-commerce websites.

Database is everywhere



- Student/Course records
- Banking transactions
- Library book records (http://www.hkbu.edu.hk/lib)
- Online bookshop (http://www.amazon.com/)

Example database



Courses

cid	cname	credit
COMP2007	Program	3
COMP4017	Security	3
COMP2016	Database	3

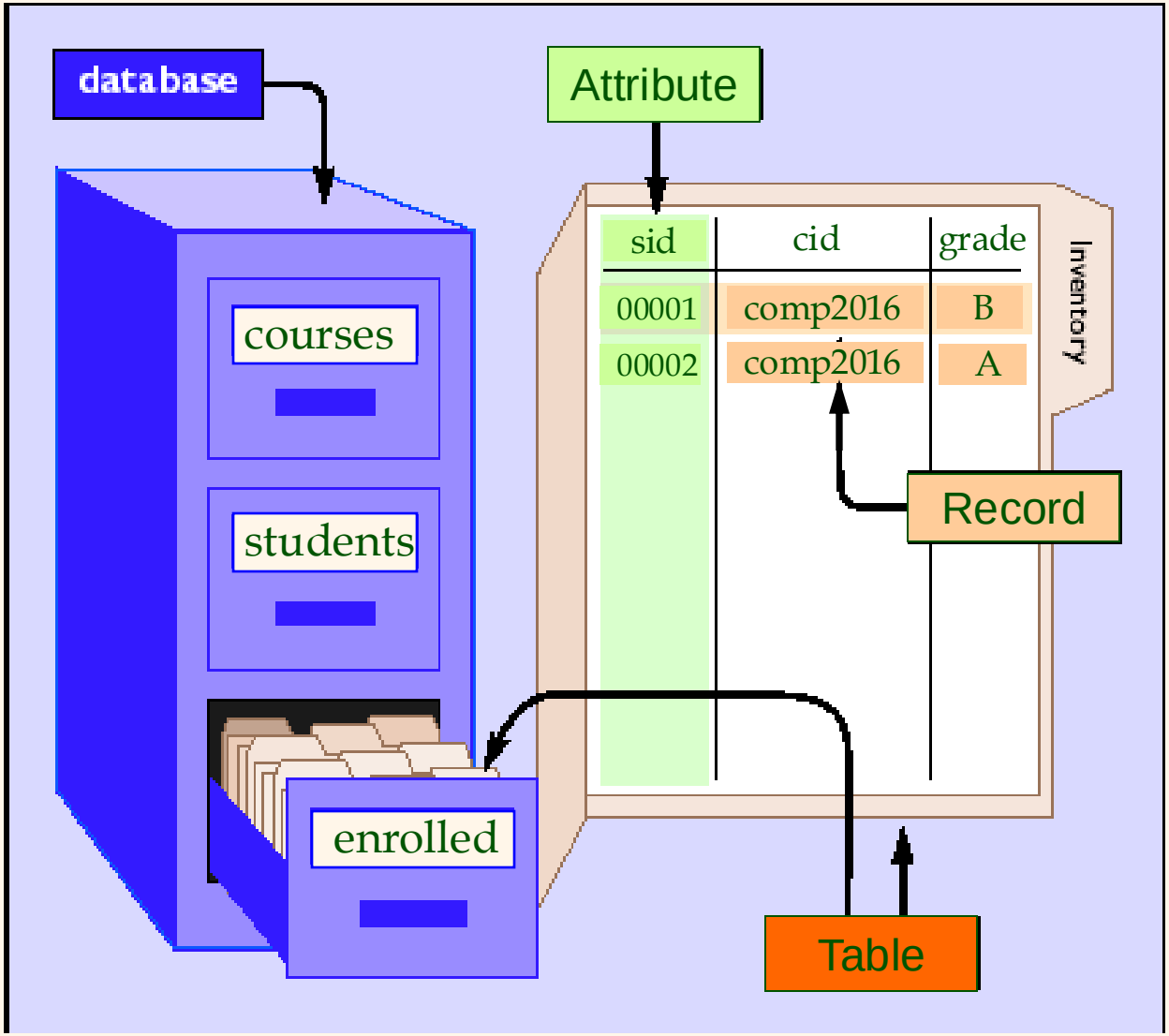
Students

sid	name	dept	gpa
00001	Jones	cs	3.4
00002	Joe	cs	3.2
00003	Smith	math	3.8

Enrolled

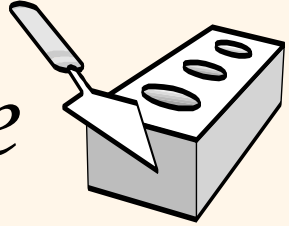
sid	cid	grade
00001	COMP2016	B
00002	COMP2016	A
00002	COMP2007	A
00003	COMP4017	B

Database concepts



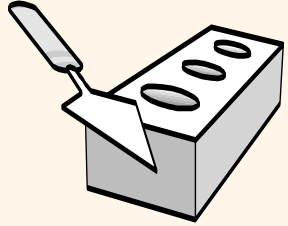
- A database consists of one or more tables
- Each table is made up of a number of records (a.k.a. tuples)
- Each record contains several attributes

Extract information from a database



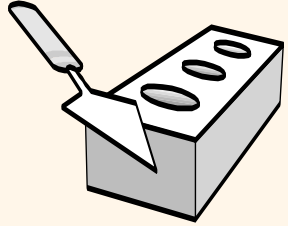
- We can ask questions for a database. For example:
  - Who has student id '1001'?
  - Which courses has Bob taken?
  - How many students have enrolled in COMP7640?
  - How many female students are in DAAI program?
- But a database doesn't speak English…

Database language (SQL)



- Who has student id ‘1001’?

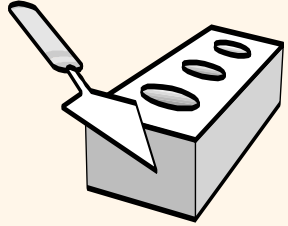
```
SELECT name
FROM   Students
WHERE  sid = ‘1001’;
```



Course Objectives

- To learn how to ***manage*** such a database. For example,
  - Design a database.
  - Query a database.
  - Disk and memory management
  - Access methods and indexing
  - Query evaluation and optimization
  - Concurrency control and crash recovery

References



- A. Silberschatz, H. F. Korth, and S. Sudarshan, *Database System Concepts*, McGraw-Hill, 2019.

- R. Ramakrishnan and J. Gehrke, *Database Management Systems, 3rd Edition (ISBN 0-07-115110-9)*, McGraw-Hill, 2003.