CS442
Database Management Systems

Fall 2022
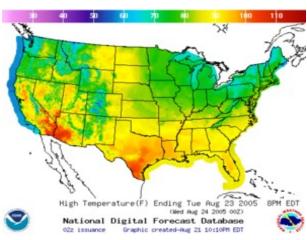
Instructor: Prof. Wendy Wang

What Is Data?

What Is a Database?

What Is a Database Management System (DBMS)?

The Origin of DBMS

- **First DBMS in history**
  - Integrated Data Stores (IDS) by Charles Bachmen, 1960.
  - Dr. Bachmen received the Turing award for IDS creation in 1973.

Database Systems: Then

- **IBM 7070: the first data processing system**

- **Built in 1960**

- **The machine is of 27 TIPS**

- **Price: $813,000 (in 1960!).**

Database Systems: Today

Show people who are:

- ☐ Male
- ☑ Female
- ☐ Unknown gender
- ☑ Single
- ☐ In a Relationship
- ☐ Married
- ☐ Open Marriage
- ☐ Unknown status

People matches: 1 - 20 of 203   1 **2 3 4 5 6 7** ... **11 Next**

Wendy

Brenda

Melissa

From:	**Jonathan** You are connected to Jonathan through: You ⇔ Melissa ⇔ Jonathan
Date:	September 1, 2002 3:30 PM
Subject:	Hello!
Message:	Hi Cindy, I'm a friend of Melissa's, and I like your profile. What kind of teaching do you do? Maybe we can play tennis sometime... Jonathan

7

Databases you may use

Databases you may use...without awareness (1/3)

- **Example 1: DB in your pocket**
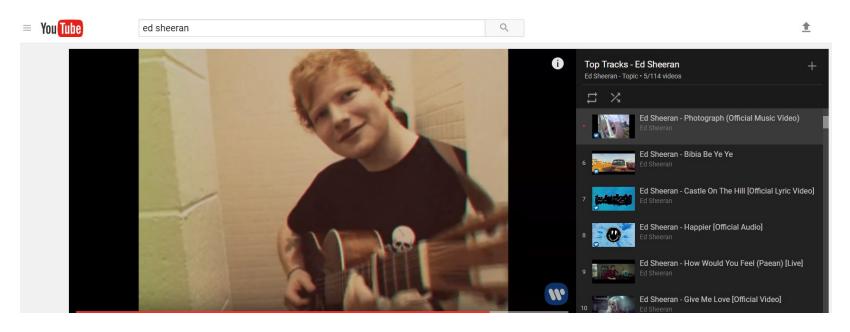
Databases you may use...without awareness (2/3)

- **Example 2: DB of social networks**

Databases you may use...without awareness (3/3)

- **Example 3: Multimedia DB**

Data Everywhere!

- **Many real-world examples of database systems:**

  - A web course registration site ;

  - A client tracking application for social service organisations;

  - A medical record system for a health care facility;

  - Your personal address book in your e-mail client;

  - ...

What's Now?

BIG DATA!!!

How much data?

- The amount of data in the world was estimated to be **44 zettabytes** at the dawn of 2020.

- By 2025, the amount of data generated each day is expected to reach **463 exabytes** globally.

- Google, Facebook, Microsoft, and Amazon store at least **1,200 petabytes** of information.

640K ought to be enough for anybody.

Why Does Big Data Matter?

- It's not the size of your data that matters!

- IT'S HOW YOU USE IT!

"Your recent Amazon purchases, Tweet score and location history makes you 23.5% welcome here."

Healthcare: COVID-19 Data Analytics

Understanding the COVID-19 Pandemic as a Big Data Analytics Issue

Big data analytics techniques are well-suited for tracking and controlling the spread of COVID-19 around the world.

Data Platform Tracks Hospital Bed Capacity During COVID-19 Outbreak

Definitive Healthcare and Esri have launched a data platform that tracks US hospital bed capacity during the COVID-19 pandemic.

Do databases always make life better?

(Source: BBC News Online, July 1, 2003)

•"Players could finally sign up for the Star Wars Galaxies game last week as Sony opened up registration to the public."

•"Once players got in to the game they found that the game servers were offline because of database problems."

•"Some players spent hours tuning their in-game characters only to find that crashes deleted all their hard work."

Do databases always make life better?

Federal Database Problems Repeatedly Delay Already Slow Waits At Illinois Driver's Services Facilities

August 6, 2020

18

Is Database Management a Simple Job?

- Consider Facebook!
- What are its database needs?

Is Database Management a Simple Job?

- Consider Facebook!
- What are its database needs?
  - Store data
  - Manipulate and present data
    - Users' personal information
    - Friendship relationships
    - Management of messages
    - Advertisements ($$$)
    - Security and privacy
    - ...
- NOT an easy job!

Is Database Management a Simple Job?

- Suppose you are building a course enrollment database system for Stevens.

- You'll need to figure out:

Is Database Management a Simple Job?

- Suppose you are building a course enrollment database system for Stevens.
- You'll need to figure out:
  - How do we store the data? (file organization, etc.)
  - How do we query the data? (write programs…)
  - Make sure that updates don't mess things up?
  - Provide different views on the data? (registrar versus students)
  - How do we deal with crashes?
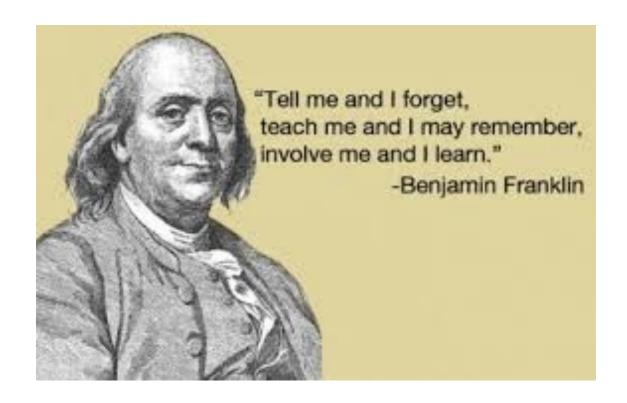
What We Will Learn in CS442?

Roughly speaking:

1. How to conceptually model the concepts in a database (DB)?

2. How to logically model concepts in a database?

3. How to do DB design and make the designed DB efficient?

4. How to write DB queries?

5. How to create DB applications?

And enjoy!

CS442 Teaching Philosophy

"Tell me and I forget,
teach me and I may remember,
involve me and I learn."
-Benjamin Franklin

EXPECT LOTS OF INVOLVEMENTS IN CS442!

Who will help you?

- **Instructor:**   **Prof. Wendy Wang (**hwang4@stevens.edu**)**

- **Teaching Assistants (TA/CAs):**
  - Aughdon Breslin <abreslin@stevens.edu>,
  - Sebastian Churion <schurion@stevens.edu>,
  - Alexander Rubino <arubino@stevens.edu>,
  - Dennis Salmanowitz <dsalmano@stevens.edu>,
  - Grace Mattern <gmattern@stevens.edu>

The list of TAs and our contact information are also included in "Course Syllabus" in Canvas

Office Hour

– Instructor Office Hour (weekly): 2pm-3pm, Mondays
(Zoom link: https://stevens.zoom.us/j/99863886412)

– TA Office Hour (weekly):
- Aughdon Breslin: 12pm - 1pm Mondays
- Sebastian Churion: 2pm - 3pm Mondays
- Alexander Rubino: 5pm - 6pm Tuesdays
- Dennis Salmanowitz: 1pm - 2pm Mondays
- Grace Mattern: 4pm - 5pm Tuesdays

The schedule and location of office hours (both in-person and zoom links) are included in "Course Syllabus" in Canvas

How to reach out to the instructor?

- I will be available via email (hwang4@stevens) and will respond as soon as I am available (generally within 24 hours).

- When emailing me, please place in the subject line the course number/section and the topic of the email (i.e. CS442 – Assignment 2 Question).

How to communicate with TAs/CAs?

- TAs will be available via emails too.
- When emailing TAs, please place in the subject line the course number/section and the topic of the email (i.e. CS442 – Assignment 2 Question).
- When submitting your assignments, an assignment file should be appended by your username, such as "assignment1_userID.doc". This may make it easier for TAs to manage assignment files.
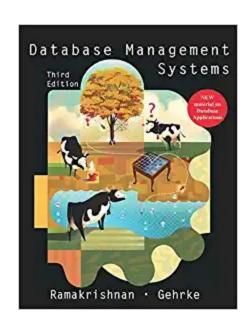
Course Prerequisites

- **Required: CS 385 Data Structures and Algorithms II**

Course Materials

- **Recommended:**
  - Database Management Systems, 3rd Edition, Raghu Ramakrishnan and Johannes Gehrke, McGraw-Hill, 2002.
  - What if you have already bought the 2nd edition?
    - Don't despair! You can make do with it.

Lectures & Notes

- Lectures may not follow text closely

- You are responsible for the textbook, lectures, and any additional reading that may be assigned

- Lecture notes will be available <u>before class</u> at Canvas

Taking Notes in Class

- Some slides have an *Notebook* icon

at the top corner (see the top corner for the example).

- Typically these slides are for exercises without solutions.

- This means that you need to take notes for the solutions!

Grading Rubric

- **Attendance (5%)**
- **In-class quizzes (5%)**
- **In-class lab (10%)**
- **Homework assignments (30%)**
- **Midterm exam (20%)**
- **Final exam (30%)**

Attendance (5% of Final Grade)

- All students need to fill in the sign-in attendance in class

- Excused attendance
  - Medical reasons, family issues, …
  - Email instructor with evidence (doctor's notes, family members' email, etc.) before the class you will be absent

- Three "FREE treats" besides excused attendance
  - Each student can be absent for up to THREE lectures

Assignments (30% of Final Grade)

- **Assignment hand-in policy**
  - Submit via Canvas

- **Late penalty**
  - 20% of the total points will be deducted for assignments received 1-6 days late;
  - Assignments received more than 1 week late will receive 0 points.

Missing Assignments

- For those assignments missing in Canvas but students claimed they submitted on time:
    - Still considered as late submission;
    - File timestamps are not accepted as the evidence that the assignments were finished in time;
    - Always check if your assignment was submitted successfully.

Make-up Exams and Extra Credits

- **Make-up exam**
  - A make-up for the quizzes will be granted only if
    - there is serious illness (proven by doctor's notes) or similarly important reason for missing the exam.
  - A make-up for midterm/final exam will be granted only if
    - The instructor is notified BEFORE the exam; and
    - There is serious illness (proven by doctor's notes) or similarly important reason for missing the exam.
- **Extra credit**
  - For fairness to all students there will be NO individual extra credit work.

Policies on Cheating

- **It is cheating to**
  - Copy a solution/material from another student or some other resource without specifying it;
  - Give away a solution/material

- **Zero Tolerance**
  - Any case of cheating will be reported to the honor board of Stevens.

Grade Complaints

- Any complaint regarding a grade must be presented no later than 7 weekdays following the pronunciation of grades of respective assignments and midterm.

More Information of Course Policy in Canvas

- **Check "Course syllabus" in "Course Content"**
  - Email addresses of the instructor and TAs;
  - Office hours and zoom links;
  - Course schedule;
  - Exam procedures;
  - ...

Finally, what about you?

- Finish the 5-minute survey available in Canvas under "Quizzes".

- The survey will be closed at midnight of Thursday (9/1).