Lecture 1:

- Survey
- Course Overview
- Basic DB Concepts and Terms

Homework:

- GitHub account
- Piazza
- Quiz 1 Tues
Course Logistics

1. Course webpage: www.cs.gonzaga.edu/bowers/courses/cpsc321

2. Piazza: for Q&A, announcements (see invite)

3. GitHub: for submitting homework

4. Canvas: for tracking points

5. Office Hours: Tu/Th 3:30–4:30, Wed 1:30–3:30

6. Graders: Kevin Dang () and Vincent Do ()

7. HW Late Policy: Up to 1-week late, 20% penalty (email grader)

8. Grading: Out of 700 possible points:
   - 8 homework assignments at 35 points each 280 points
   - Final project 70 points
   - 10 quizzes at 10 points each 100 points
   - 2 exams at 100 points each 200 points
   - Attendance (approximately 25 lectures) 50 points

Must score at least 60% on homework (assignments + final project) and 60% on tests (quizzes + exams)

See webpage for syllabus, weekly schedule, homework assignments, etc.
Course Overview

Main Topics:

- Relational model
- Basic, intermediate, and (some) advanced SQL (in MySQL)
- Programming language and application support
- Logical database design (ER, normalization)
- Physical database design (indexes)
- Transaction (concurrency) support

Emphasis on application development as opposed to DB system implementation

Mix of basic theory and hands-on practice

Expectations:

- Engage and participate in class (including doing your own work)
- Start assignments early, give yourself enough time to succeed
- Assume you have everything you need (ask when in doubt)
- Study for quizzes, exams, etc
- Come to office hours and frequently check piazza

(Some) collaboration is encouraged ... e.g., discussing assignments

- but avoid plagiarism & other related issues — e.g., no code sharing
Basic Concepts and Terms

1. **Persistence**: store data beyond (running of) application that created it

2. **Database**: an organized (structured) data store for information
   - the organization makes it easier to obtain info and answer questions
   - e.g., find pictures taken in Spokane from 2020–2022

3. **Database Management System** (DBMS): software for managing databases

4. A DBMS helps users to:
   - create an organization for data ... called a “schema”
   - update what is stored by adding, removing, modifying data
   - retrieve data ... using a “query language”
   - manage who has permission to access data

5. **Relational DBMSs and SQL**
   - employs relational data model, storing data in tables of rows and columns
   - SQL is the standard query language used by (relational) DBMSs

---

**Check-In:**

- What does “persistence” mean in terms of a database systems?
- What are the four main ways a DBMS helps users?
- What is mean by a “Relational DBMS”?
6. **SQL Command (Sub) Languages**

- Data Definition Language (DDL) ... CREATE, ALTER, DROP
- Data Manipulation Language (DML) ... INSERT, UPDATE, DELETE
- Data Query Language (DQL) ... SELECT, WITH
- Data Control Language (DCL) ... GRANT, REVOKE (accounts, privileges)

*Note:* other data models exist as well

- e.g., document, tree, graph, key-value, time-series models

7. **Transaction Processing** (aka “OLTP”)

- supports the “daily operations” of a business (in real time)
- many insert, update, delete, and simple queries
- fast response times and many concurrent sessions

8. **Data Analytics** (aka “OLAP”)

- supports “decision making” within an organization (business)
- often more complicated queries over large portion of database
- read heavy (with fewer inserts, updates, deletes)
- today, frequently implies use of a “data warehouse”

**Check-In:**

- What are the four SQL sublanguages?
- How do transaction processing and data analytics differ?
9. **Transactional concurrency**
   - a single transaction can involve many data reads and writes (SQL commands)
   - DBMSs designed to handle many transactions at once (concurrency)

   *Note:* transaction support leads to system design and implementation challenges

10. **ACID (transaction concurrency) properties**
   - *Atomicity:* all or nothing execution of transaction
   - *Consistency:* transactions expected to preserve data integrity (constraints)
   - *Isolation:* transactions execute as if no others execute at same time
   - *Durability:* transaction effect on DB must not be lost after it completes

   *Note:* Relational DBMSs typically ACID compliant (vs “NoSQL” systems)

**Check-In:**
- What is a database “transaction”?
- What are the ACID (transaction concurrency) properties?
DBMS System Rankings

Market share based on company revenue

- Around $80 billion revenue in 2021
- About 50% from “managed cloud services”
- About 80% from top-5: Microsoft, AWS, Oracle, Google, IBM
- About 3% from non-relational vendors (primarily MongoDB)

Popularity various indicators

- Most popular (all): Oracle, MySQL, SQL Server, PostgreSQL, MongoDB
- Most popular (relational): Oracle, MySQL, SQL Server, PostgreSQL, DB2