Introduction

CSC 256, SQL Programming

Databases

- A **database** is a collection of related data organized in a way that supports easily processing the data
- A Database Management System (DBMS) is software that provides the ability to create and manipulate a database

Databases are Everywhere

- Desktop, mobile, web, embedded
- Provide reliability, scalability, performance, and flexibility
- SQL is the most common language to query and manage relational databases

DBMS versus Text

- Reliably work with terabytes of data
- Maintain data integrity
- Handle concurrency - multiple people trying to read and write the data at the same time
- Manages user permissions and access
- Replication and backups
- Programmatic interface

Relational and Non-relational

- Relational
  - data is stored in tables (relations) with minimal redundancy and data integrity is maintained
  - For example, PostgreSQL, MySQL, etc.
- Non-relational
  - NoSQL databases, key-value, graph, document, etc.
  - For example, MongoDB, Cassandra, etc.
  - Good for special purposes

Database Languages

- Data definition language (DDL): defines the data types and relationships among them
- Data manipulation language (DML): performs operations such as inserting, updating, or deleting data.
- Query language: performs information retrieval
- Some DMBSs also have a transaction control language (TCL) and an access control language (ACL)

SQL

- The Structured Query Language (SQL) combines the roles of all database languages into a single language for relational databases
- SQL is ANSI and ISO standardized
- SQL is a *declarative* programming language
- A declarative programming language describes *what* computation should be performed not *how* to compute it