Module 2 - Lecture 1

Introduction to Databases

Programs Don't Run Forever

Even infinite loops stop at some point.

A **database** is an organized collection of data that can be accessed, managed, and updated.

Who uses these things?

A **relational database** is a set of formally described tables from which data can be accessed or reassembled in many different ways without having to reorganize the database tables

- Each **table** is an entity.   (Patient)
- **Columns** are called attributes   (PatientId, Name, ...)
- **Rows** represent individual records. (134, Jeff, ...)

Patient Id	Name	D.o.B	Gender	Phone	Doctor Id
134	Jeff	4-Jul-1993	Male	7876453	01
178	David	8-Feb-1987	Male	8635467	02
198	Lisa	18-Dec-1979	Female	7498735	01
210	Frank	29-Apr-1983	Male	7943521	01
258	Rachel	8-Feb-1987	Female	8367242	02

(R)DBMS software is designed to help manage a database. Its four basic functions are (1) data definition, (2) data storage, (3) data retrieval, and (4) administration

SQL

SQL stands for structured query language and is a *declarative programming language* to retrieve and update records from a database.
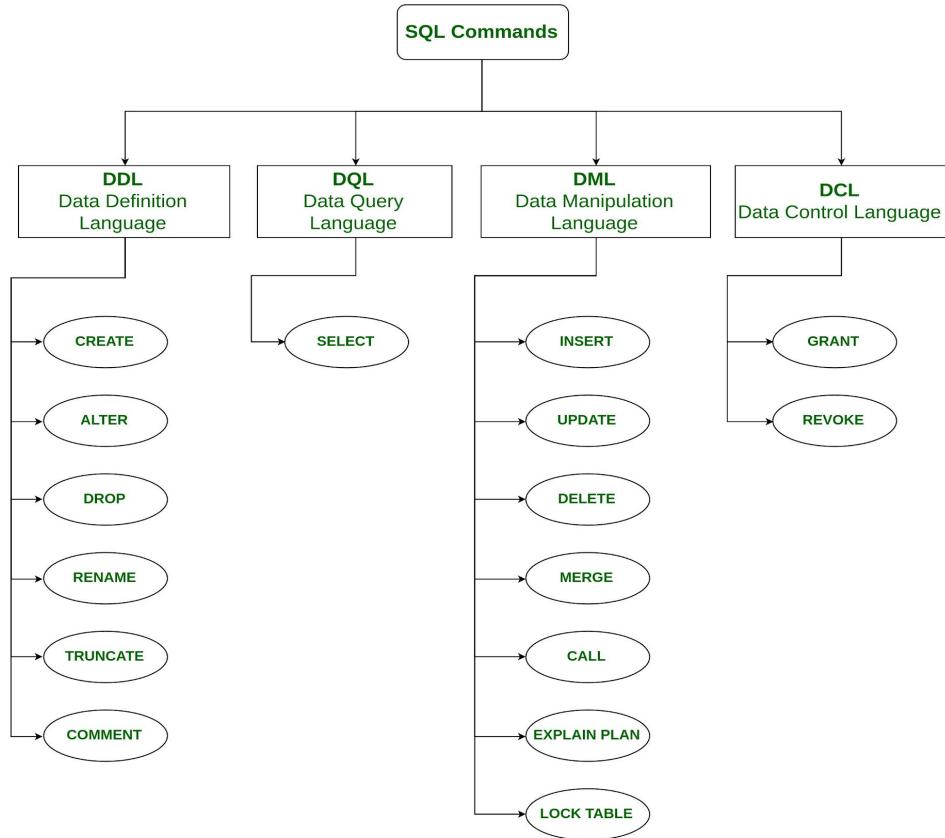
SQL consists of:

- *data definition language* to define the data structures
- *data manipulation language* to query and modify the data in a database
- *data control language* to define access to a particular database

Types of SQL Commands

```
                        ┌──────────────────┐
                        │  SQL Commands    │
                        └──────────────────┘
        ┌───────────────────┬──────────┴──────────┬───────────────────┐
┌───────────────┐  ┌───────────────┐  ┌───────────────────┐  ┌────────────────────┐
│     DDL       │  │     DQL       │  │       DML         │  │        DCL         │
│ Data Definition│  │ Data Query   │  │ Data Manipulation │  │ Data Control Language│
│   Language    │  │  Language    │  │    Language        │  │                    │
└───────────────┘  └───────────────┘  └───────────────────┘  └────────────────────┘
```

DDL	DQL	DML	DCL
CREATE	SELECT	INSERT	GRANT
ALTER		UPDATE	REVOKE
DROP		DELETE	
RENAME		MERGE	
TRUNCATE		CALL	
COMMENT		EXPLAIN PLAN	
		LOCK TABLE	

GG

Let's focus on...

SELECT

The **SELECT clause** indicates what columns that you want to get from a database table.

The **FROM clause** indicates which table(s) to retrieve the data from.

```
SELECT column1, … ,[column-n] FROM table;
```

The **DISTINCT clause** indicates that duplicate values should not be included.

The **AS clause** establishes an alias for a particular column name

```
SELECT DISTINCT column AS blah FROM table;
```

The **WHERE clause** is used to filter the result set using one or more criteria rules.

- Conditional clauses in the WHERE clause can include
  - =, <>, !=, >, >=, <, <=
  - IN(values), NOT IN(values)
  - BETWEEN value AND value
  - IS NULL, IS NOT NULL
  - LIKE (with wildcard characters)
- Multiple conditions use AND and OR

QUESTIONS?