

# Database Management Systems

SQL (II)

# Data Manipulation

- Case Insensitive Language (except in literal strings)
- Uses BAG Semantics — allow duplicates unless explicitly removed
- Correctness
  - Retrieves the qualified data
  - Retrieves all the qualified data
  - Retrieves nothing but the qualified data
  - Works correctly on ALL possible database instances
- Input of SQL query: relations
- output of SQL query: a relation

# SQL Keywords

- **SELECT:** specifies which columns are to appear in the output
- **FROM:** specifies the table or tables to be used
- **WHERE:** filters the rows subject to some condition expressed as boolean expression
- **GROUP BY:** forms groups of rows with the same column value
- **HAVING:** filters the groups subject to some condition expressed as boolean expression
- **ORDER BY:** specifies the order of the output

# Sample Database Schema

- Departments(did, dname, managerId)
- Emps(eid, firstname, lastname, salary, workdept, hireDate)
- Projects(pno, title, respEid, respDid)
- WorksOn(eid, pno, sdate, edate)

# A Simple Example

- Query:  
Select firstname, lastname, hiredate  
From Emps  
Where salary >= 60000;

- Result:

FIRSTNAME	LASTNAME	HIREDATE
Anne	Apland	10-SEP-08
Cathy	Carren	12-NOV-08
Dave	Denn	20-MAR-09
Ethan	Edin	15-APR-12

# Select Block

- SQL functions
- String concatenation
- Renaming columns
- Eliminating duplications (distinct)

# Where Condition

- Optional
- Boolean expression
  - Logical operators: AND, OR, NOT
  - Comparison operators: >, <, >=, <=, =, <>
- Eliminating ambiguity by prefix the column name with table name
- String pattern match
  - operator: like
  - wild card matching symbol: \_ and %

# Concept of Join

- cross product
- join condition
- foreign key join
- natural join
- difference between join condition and some of the where condition
- places to put the join condition



# From List

- Can't appear alone
- relation vs relation variable
- all relation variables are distinct, even if they represent the data from the same relation
- re-name a relation
  - to simplify query writing
  - to eliminate the ambiguity
- sub queries in from list