Chapter #2:

**The Relational Database Model**

* **Logical View of Data:**
* **Relational Database:**
  + - The focus is on logical representation rather than physical representation.
    - User is only interested in the logical representation (what data is being collected).
    - RDBMS is concerned with “How data is collected”.
    - Advantages of using Tables (Relations):
      * Structural and data independence
      * Related records are stored in independent table.
      * Logical simplicity
    - It allows more effective design options.
* **Entities and Attributes:**
  + - Entity is a noun.
    - It could be a person, place, event, or thing about which data is collected.
    - Entity cannot be any object. Only those objects are entities, about which we collect data.
    - Attributes are characteristics of the entity, which describes that entity.
* **Tables:**
  + - A table only holds related entities or entity set. (Entities with common characteristics are Related)
    - Mathematical name of Table -> Relation
    - A table contains rows and columns.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

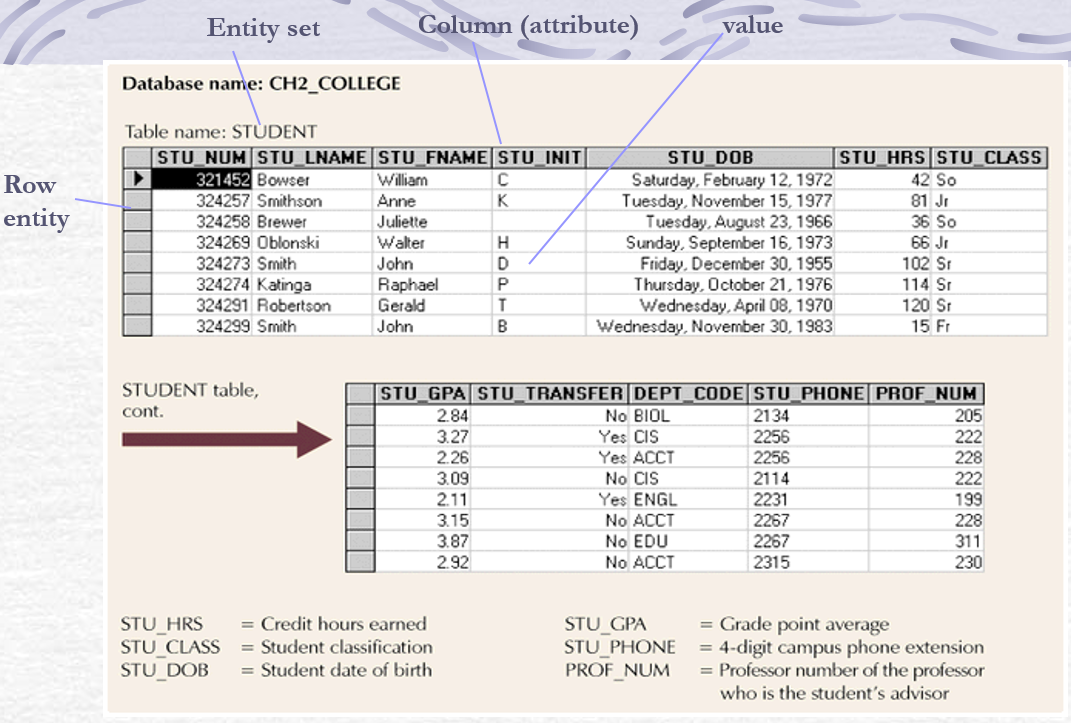
* **Table Characteristics:** (IMPORTANT)
  + A table is a 2D structure with rows and columns.
  + Rows (tuples) represent single entity.
  + Columns represent attributes.
  + The intersection of Row & column represents single value.
  + Tables must have a uniquely identifying attribute (UID) or primary key for each row.

{Primary key: attribute and a combination of combined attributes that uniquely identify any given row}

* + Column values have same data format/ data types.
    - Data types are:
      * Number
      * Character
      * Date
      * Logical
  + Each column has range of values called attribute domain.

{To implement range of values, we apply constraint}

* + - **For example**; if we make a column of “CGPA” in student data table, the values of CGPA would range between ‘0’ and ‘4’.
  + Order of the rows and columns is immaterial to the DBMS (means Order of either rows or columns doesn’t matter in DBMS)



* **RELATIONAL DATABASE KEYS:** (MOST IMPORTANT CONCEPT)

**“Sample Table” for understanding concepts:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Student ID (SID) | Student Name (sName) | CNIC number  (NIC) | Address | CGPA |

1. **SUPER KEY:**

* An attribute (or combination of attributes) that uniquely identifies each entity in a table.
* We can also apply conditions/ constraints on combinations of attributes.
* **For example:**

From sample table,

1. SID
2. sName + SID
3. NIC
4. NIC + SID
5. Address + NIC
6. sName + Address
7. sName + Address + SID
8. SID + NIC + CGPA + Address
9. sName + Address + CGPA

.

.

. Etc.

**Note:** As there can be multiple combinations of attributes which can be super keys, so it is complex.

1. **CANDIDATE KEY:**

* Candidate key is based on super key.
* A minimal super key.
* It is a solution of super key.
* A super key that does not contain subset of attributes that is itself a super key.
* Subset of attributes (that is itself a super key) could also be a combination of attributes.
* From sample table, considering same super key examples.

1. SID
2. ×

9. sName + Address + CGPA

Because sName + Address is also a super key as combination.

1. NIC
2. ×
3. ×
4. sName + Address
5. ×
6. ×
7. ×

* **Examples:**

1. SID
2. NIC
3. sName + Address
4. **PRIMARY KEY:**

* A selected candidate key is called Primary Key.
* A candidate key which is selected to uniquely identify all other column values in a given row.
* It is a Unique Identifier.
* It is only one attribute or column.
* It cannot contain NULL values.
* **For example:**

From sample table,

1. SID is a primary key.
2. **SECONDARY KEY:**

* An attribute (or combination of attributes) that narrow down our searching criteria.
* It is strictly used for data retrieval purposes.
* It does not refer to primary key.
* It may or may not be a super key.

1. **FOREIGN KEY:**

* An attribute which builds relationship with primary key of targeted table.
* An attribute of one table whose values must match with the primary key values of targeted table (except NULL values)
* **INTEGRITY RULES:**

1. **Entity integrity:**
   * + It applies on primary key.
     + The requirement on Entity Integrity:

{Ensure that all entities are unique, no NULL or repeated (duplicate) values}

1. **Referential integrity:**

* It applies on foreign key.
* Foreign key values must match primary key except NULL values.
* Primary key may or may not match with foreign key, but foreign key must match.
* One value of Primary key can associate with multiple values of foreign key.
* There exist Parent-Child Relationship (1: M) between Primary and Foreign key.
* **For example:**

Consider the “***deptNo***” column/attribute of EMP and DEPT tables.

-> ***deptNo*** from EMP acts as foreign key.

-> ***deptNo***from DEPT acts as primary key.

DEPT table

EMP table

**PARENT**

**CHILD**

M : 1

10

20

30

10

20

30

55

NULL

* **Simple Relational Database:**

