DBMS

1)What is a primary key?

A primary key is a combination of fields which uniquely specify a row. This is a special kind of unique key, and it has implicit NOT NULL constraint. It means, Primary key values cannot be NULL.

2)What is a foreign key?

A foreign key is one table which can be related to the primary key of another table. Relationship needs to be created between two tables by referencing foreign key with the primary key of another table.

3)What is a join?

This is a keyword used to query data from more tables based on the relationship between the fields of the tables. Keys play a major role when JOINs are used.

What are the types of join and explain each?

There are various types of join which can be used to retrieve data and it depends on the relationship between tables.

- **Inner Join.**

Inner join return rows when there is at least one match of rows between the tables.

- **Right Join.**

Right join return rows which are common between the tables and all rows of Right hand side table. Simply, it returns all the rows from the right hand side table even though there are no matches in the left hand side table.

- **Left Join.**

Left join return rows which are common between the tables and all rows of Left hand side table. Simply, it returns all the rows from Left hand side table even though there are no matches in the Right hand side table.

Full Join.

Full join return rows when there are matching rows in any one of the tables. This means, it returns all the rows from the left hand side table and all the rows from the right hand side table.

What is normalization?

Normalization is the process of minimizing redundancy and dependency by organizing fields and table of a database. The main aim of Normalization is to add, delete or modify field that can be made in a single table.

What is Denormalization.

DeNormalization is a technique used to access the data from higher to lower normal forms of database. It is also process of introducing redundancy into a table by incorporating data from the related tables.

What is a View?

A view is a virtual table which consists of a subset of data contained in a table. Views are not virtually present, and it takes less space to store. View can have data of one or more tables combined, and it is depending on the relationship.

What is an Index?

An index is performance tuning method of allowing faster retrieval of records from the table. An index creates an entry for each value and it will be faster to retrieve data.

What is a Cursor?

A database Cursor is a control which enables traversal over the rows or records in the table. This can be viewed as a pointer to one row in a set of rows. Cursor is very much useful for traversing such as retrieval, addition and removal of database records.

What is a relationship and what are they?

Database Relationship is defined as the connection between the tables in a database. There are various data basing relationships, and they are as follows:.

- One to One Relationship.
- One to Many Relationship.
- Many to One Relationship.
- Self-Referencing Relationship.

types of subquery:

There are two types of subquery – Correlated and Non-Correlated.

A correlated subquery cannot be considered as independent query, but it can refer the column in a table listed in the FROM the list of the main query.

A Non-Correlated sub query can be considered as independent query and the output of subquery are substituted in the main query.

A noncorrelated (simple) subquery obtains its results independently of its containing (outer) statement.A query's `WHERE` and `HAVING` clauses can specify noncorrelated subqueries if the subquery resolves to a single row,

A correlated subquery requires values from its outer query in order to execute.

What is a stored procedure?

Stored Procedure is a function consists of many SQL statement to access the database system. Several SQL statements are consolidated into a stored procedure and execute them whenever and wherever required.

What is a trigger?

A DB trigger is a code or programs that automatically execute with response to some event on a table or view in a database. Mainly, trigger helps to maintain the integrity of the database.

Example: When a new student is added to the student database, new records should be created in the related tables like Exam, Score and Attendance tables.

difference between DELETE and TRUNCATE commands?

DELETE command is used to remove rows from the table, and WHERE clause can be used for conditional set of parameters. Commit and Rollback can be performed after delete statement.

TRUNCATE removes all rows from the table. Truncate operation cannot be rolled back.

What is a constraint?

Constraint can be used to specify the limit on the data type of table. Constraint can be specified while creating or altering the table statement. Sample of constraint are.

- NOT NULL.
- CHECK.
- DEFAULT.
- UNIQUE.
- PRIMARY KEY.
- FOREIGN KEY.

What is Auto Increment?

Auto increment keyword allows the user to create a unique number to be generated when a new record is inserted into the table. AUTO INCREMENT keyword can be used in Oracle and IDENTITY keyword can be used in SQL SERVER.

Mostly this keyword can be used whenever PRIMARY KEY is used.

What is Self-Join?

Self-join is set to be query used to compare to itself. This is used to compare values in a column with other values in the same column in the same table. ALIAS ES can be used for the same table comparison.

What is Cross-Join?

Cross join defines as Cartesian product where number of rows in the first table multiplied by number of rows in the second table. If suppose, WHERE clause is used in cross join then the query will work like an INNER JOIN.

Advantages and Disadvantages of Stored Procedure?

Stored procedure can be used as a modular programming – means create once, store and call for several times whenever required. This supports faster execution instead of executing multiple queries. This reduces network traffic and provides better security to the data.

Disadvantage is that it can be executed only in the Database and utilizes more memory in the database server.

What is Online Transaction Processing (OLTP)?

Online Transaction Processing (OLTP) manages transaction based applications which can be used for data entry, data retrieval and data processing. OLTP makes data management simple and efficient. Unlike OLAP systems goal of OLTP systems is serving real-time transactions.

Example – Bank Transactions on a daily basis.

What is CLAUSE?

SQL clause is defined to limit the result set by providing condition to the query. This usually filters some rows from the whole set of records.

Example – Query that has WHERE condition

Query that has HAVING condition.

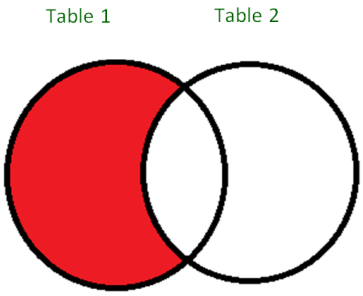
What is recursive stored procedure?

A stored procedure which calls by itself until it reaches some boundary condition. This recursive function or procedure helps programmers to use the same set of code any number of times.

What is Union, minus and Interact commands?

UNION operator is used to combine the results of two tables, and it eliminates duplicate rows from the tables.

MINUS operator is used to return rows from the first query but not from the second query. Matching records of first and second query and other rows from the first query will be displayed as a result set.



As you can see is in the above diagram, the MINUS operator will return only those rows which are present in the result set from Table1 and not present in the result set of Table2.

Basic Syntax:

```
SELECT column1 , column2 , ... columnN

FROM table_name

WHERE condition

MINUS

SELECT column1 , column2 , ... columnN

FROM table_name

WHERE condition;
```

columnN: column1, column2.. are the name of columns of the table.

Important Points:

- The WHERE clause is optional in the above query.
- The number of columns in both SELECT statements must be same.
- The data type of corresponding columns of both SELECT statement must be same.

INTERSECT operator is used to return rows returned by both the queries.

difference between TRUNCATE and DROP statements?

TRUNCATE removes all the rows from the table, and it cannot be rolled back. DROP command removes a table from the database and operation cannot be rolled back.

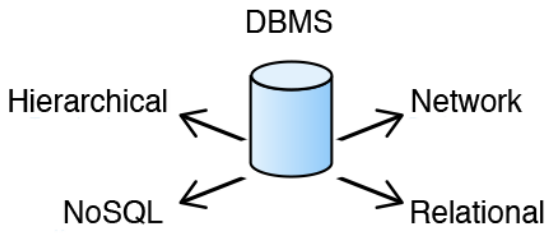
. What are aggregate and scalar functions?

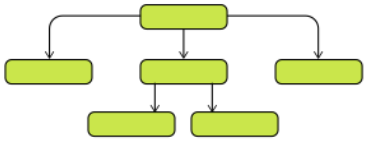
Aggregate functions are used to evaluate mathematical calculation and return single values. This can be calculated from the columns in a table. Scalar functions return a single value based on the input value.

Example -.

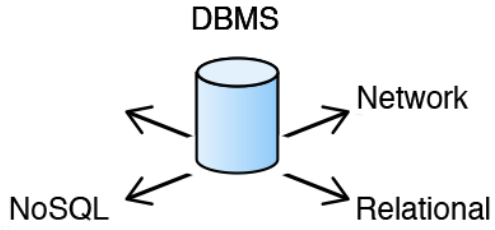
Aggregate – max(), count - Calculated with respect to numeric.

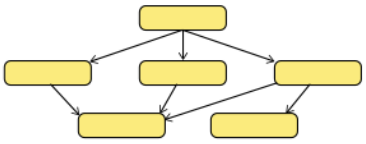
Scalar – UCASE(), NOW() – Calculated with respect to strings.



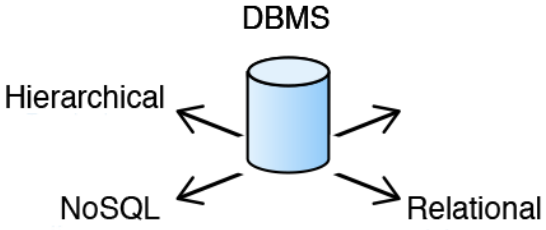


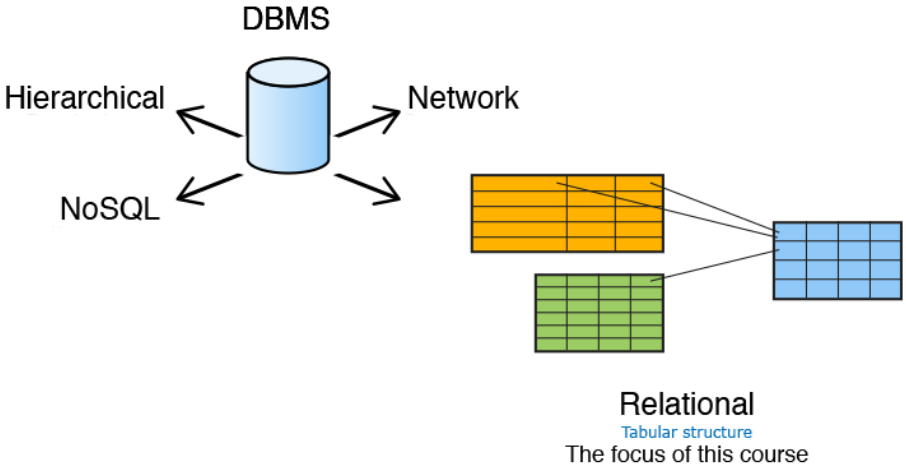
Hierarchical
Tree structure

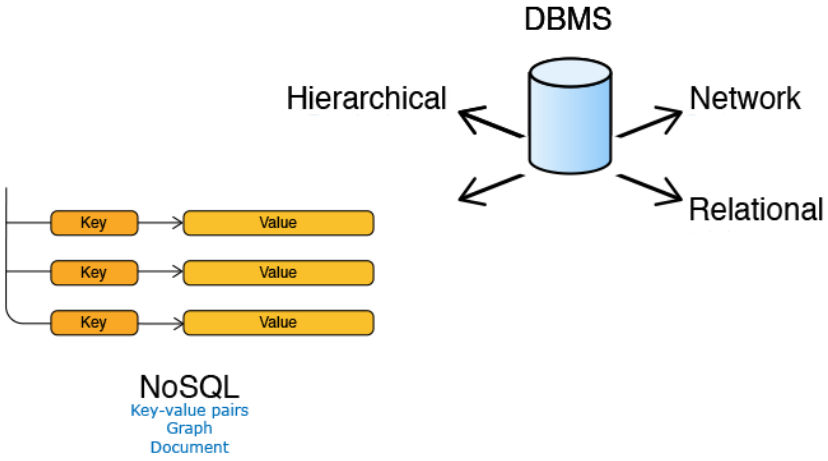




Network
Graph structure





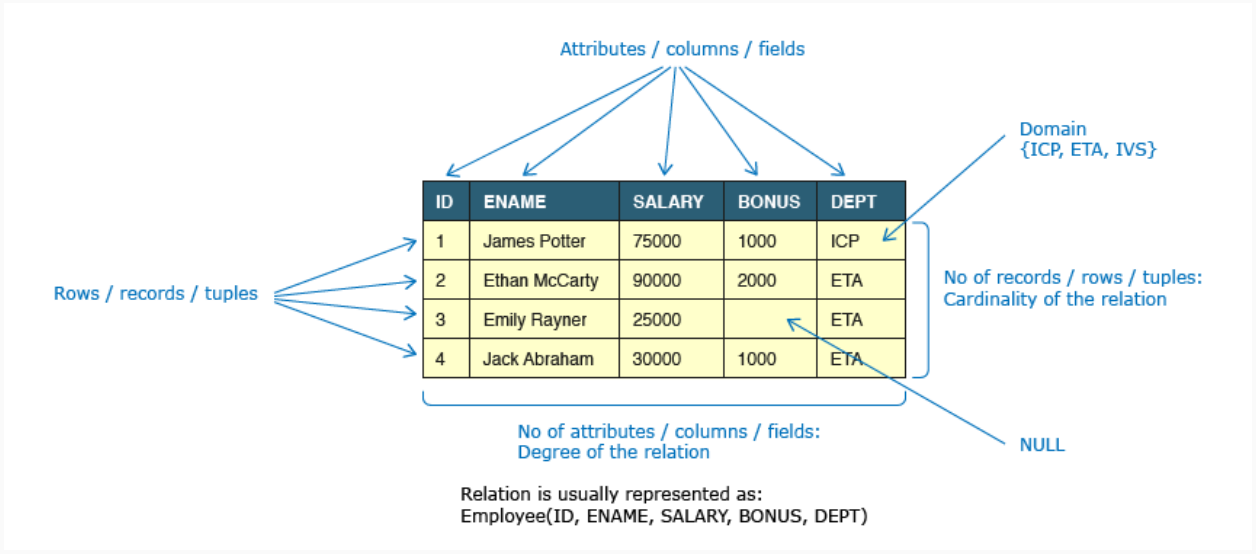


Cardinality of relation is the number of rows it contains

Degree of relation is the number of attributes it contains

A **domain** is the set of allowable values for one or more attributes.

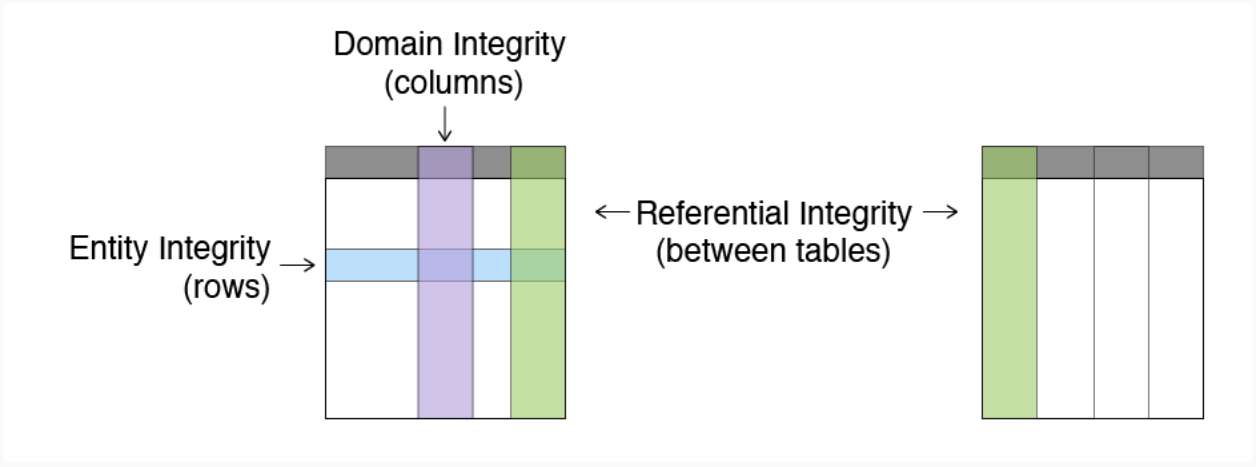
A collection of relations with distinct relation names is called as Relational Model.



ID	ENAME	SALARY	BONUS	DEPT
1	James Potter	75000	1000	ICP
2	Ethan McCarty	90000	2000	ETA
3	Emily Rayner	25000		ETA
4	Jack Abraham	30000	1000	ETA

Data integrity refers to maintaining and assuring the accuracy and consistency of data over its entire life-cycle.

Integrity Types	Definition	Enforced Through
Entity Integrity	Each table must have a column or a set of columns through which we can uniquely identify a row. These column(s) cannot have empty (null) values.	PRIMARY KEY
Domain Integrity	All attributes in a table must have a defined domain i.e. a finite set of values which have to be used. When we assign a data type to a column we limit the values that it can contain. In addition we can also have value restriction as per business rules e.g. Gender must be M or F.	DATA TYPES, CHECK CONSTRAINT
Referential Integrity	Every value of a column in a table must exist as a value of another column in a different (or the same) table.	FOREIGN KEY



Candidate Key is a minimal set of columns/attributes that can be used to uniquely identify a single row/tuple in a relation.

A **Primary Key** is the candidate key that is selected to uniquely identify a tuple in a relation.

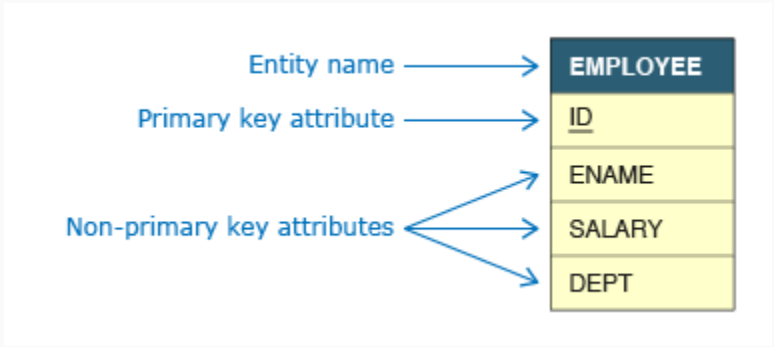
When two or more columns together identify the unique row then it's referred to as **Composite Primary Key**

A Foreign Key is a set of one or more columns in the child table whose values are required to match with corresponding columns in the parent table.

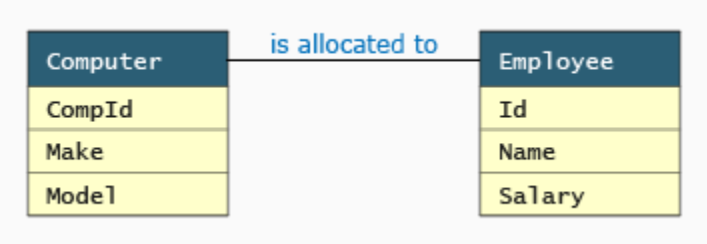
ER model is a graphical representation of entities and their relationships which helps in understanding data independent of the actual database implementation.

Term	Definition	Examples
Entity	Real world objects which have an independent existence and about which we intend to collect data.	Employee, Computer
Attribute	A property that describes an entity.	Name, Salary

A sample ER Diagram representing the Employee entity along with its attributes is presented below:

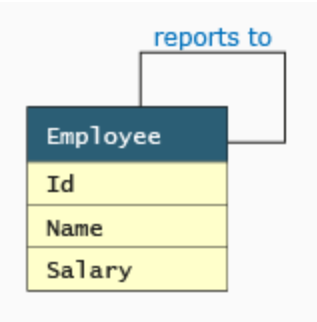


Relationships are associations of one entity with another entity through a foreign key. Each relationship has a name e.g. a Computer is allocated to an Employee.



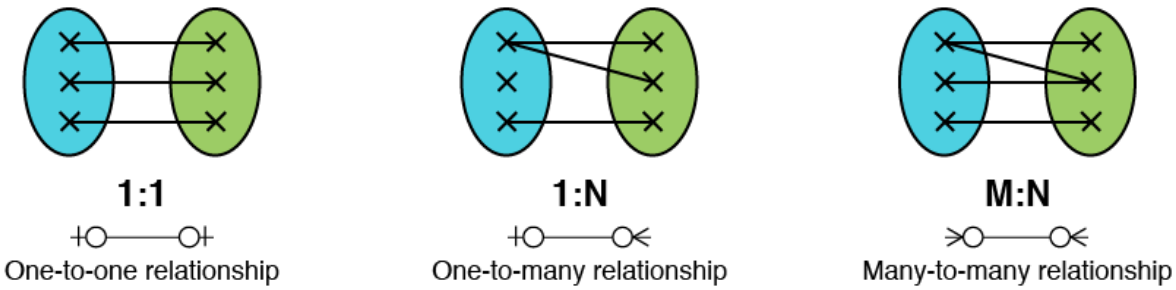
There can be more than one relationship between entities, e.g. an Employee works in a Department while the head of the department (also an employee) manages a Department.

A relationship can also exist between instances of the same entity, e.g. an Employee reports to a manager (also an Employee).

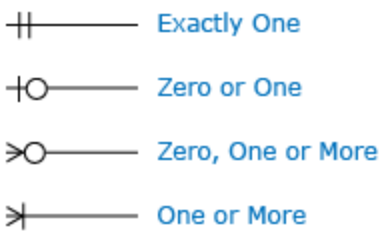


Cardinality of relationship is the number of instances in one entity which is associated to the number of instances in another. For the relationship between Employee and Computer, it helps us answer questions like how many computers can be allocated to an employee, can computers be shared between employees, can employees exist without being allocated a computer etc. e.g. if 0 or 1 computer can be allocated to 0 or 1 employee then the cardinality of relationship between these two entities will be 1:1.

Cardinality of relationships are of three types: 1:1, 1:N and M:N.



Crow foot notation is one of the ways to represent cardinality of relationship in an ER Model. The notation comprises of four symbols and one of them needs to be used for each entity in a relationship.

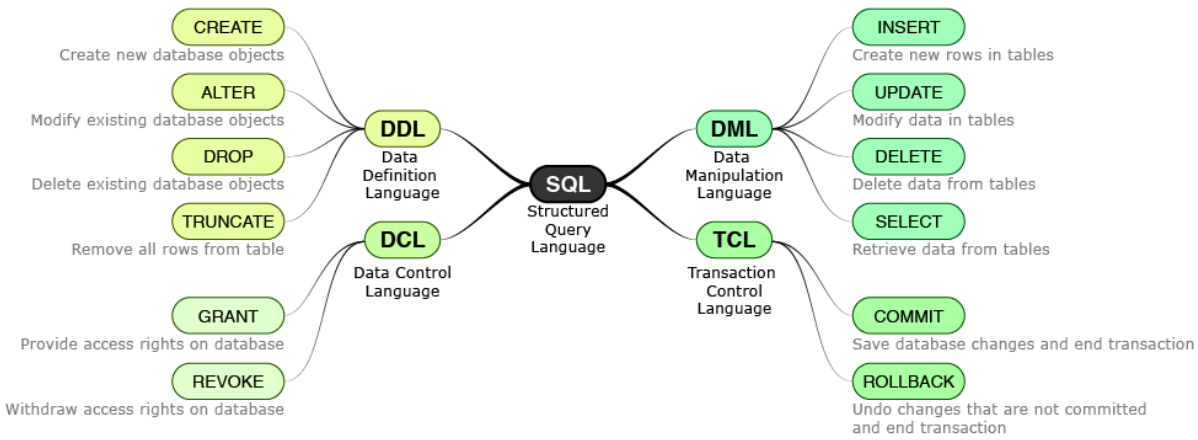


Converting ER diagram into tables:

In **one to one relationship**,either place a primary key attribute of one table as a foreign key attribute in the another table,after placing the foreign key attribute,insert the relationship attribute in the table where the foriegn key attribute in placed

In **one to many relationship**,place the primary key attribute of one table having one relation with the table having many relationships as a foriegn key ,insert the relationship attribute in the table where the foriegn key attribute in placed.

In **many to many relationship**,create a another table with attributes of primary key of another two table and insert the relationship attribute in the newly created table



SQL Data types :

- Character data types
- Integral data types
- Non-Integral data types
- Miscellaneous data types

SQL Operators :

- Arithmetic operators
- Comparison operators
- Logical operators

	CHAR(n)	VARCHAR2(n)
Useful for	Storing characters having pre-determined length	Storing characters whose length vary a lot
Storage size	size for n characters	size for actual no. of characters + fixed size to store length
Storage Characteristic	Trailing spaces are applied if data to be stored has smaller length than n.	Trailing spaces are not applied.
Maximum size	2000 bytes	4000 bytes
Example	A CHAR(10) field will store "Hello" as 10 bytes by appending 5 trailing spaces.	A VARCHAR2(10) field will store "Hello" as 7 bytes (assuming 2 bytes to store length).
Alternate Name	CHARACTER(n)	CHARACTER VARYING(n)

Nonintegral data types have an integer part and a fractional part. Either NUMERIC, DECIMAL or NUMBER data types can be used to store nonintegral numbers.

Comparison operators:

Operator	Symbol	Usage	Result
Equal to	=	15 = 5	false
Not equal to	<>	15 <> 5	true
Greater than	>	15 > 5	true
Greater than equal to	>=	15 >= 5	true
Less than	<	15 < 5	false
Less than equal to	<=	15 <= 5	false

Arithmetic Operators

Operator	Symbol	Usage	Result
Addition	+	15 + 5	20
Subtraction	-	15 - 5	10
Multiplication	*	15 * 5	75
Division	/	15 / 5	3

Other Comparison Operators

Operator	Symbol	Usage	Example
Range	BETWEEN <*lower limit*> AND <*upper limit*>	Matches value between a range of values (Both inclusive)	Salary BETWEEN 2500 AND 3000
List	IN (List of values)	Matches any of a list of values	Dept IN ('IVS', 'ETA', 'ICP')
String pattern matching	LIKE	Matches a character pattern	SupplierId LIKE 'S%'
NULL Test	IS NULL	Is a null value	Bonus IS NULL

Logical Operators

Operator	Symbol	Usage	Example
And	AND	Returns TRUE if both conditions are true	Salary >= 30000 AND Dept = 'ETA'
Or	OR	Returns TRUE if any one of the condition is true	Salary > 75000 OR Dept = 'ICP'
Not	NOT	Returns TRUE if following condition is false	Id NOT IN (2,3)

DROP TABLE <Table name> CASCADE CONSTRAINTS;

CASCADE CONSTRAINTS clause should be added to the DROP statement to drop all the referential integrity constraints that refer to primary and unique keys in the table.

```
CREATE TABLE Student(
   StudentId INTEGER      CONSTRAINT stud_sid_pk PRIMARY KEY,C
   FName     VARCHAR2(10) CONSTRAINT stud_fname_nn NOT NULL,C3
   LName     VARCHAR2(10) NOT NULL,C1
   Gender    CHAR(1)      CONSTRAINT stud_gender_ck CHECK(Gender IN('M', 'F')),C
   DOJ       DATE DEFAULT SYSDATE,
   ContactNo NUMBER(10)   UNIQUE,C1
   PersonId  INTEGER      CONSTRAINT stud_pid_fk REFERENCES Person(PersonId),C
   CONSTRAINT stud_name_ck  CHECK(FName <> LName)T2
)
CREATE TABLE Student(
   StudentId INTEGER,
   FName     VARCHAR2(10) CONSTRAINT stud_sname_nn NOT NULL,C3
   LName     VARCHAR2(10) NOT NULL,C13
   Gender    CHAR(1),
   DOJ       DATE DEFAULT SYSDATE,
   ContactNo NUMBER(10),
   PersonId  INTEGER,
   CONSTRAINT stud_sid_pk PRIMARY KEY(StudentId),T
   CONSTRAINT stud_gender_ck CHECK(Gender IN('M', 'F')),T
   CONSTRAINT stud_name_ck CHECK(FName <> LName),T2
   UNIQUE(ContactNo),T1
   CONSTRAINT stud_pid_fk FOREIGN KEY(PersonId) REFERENCES Person(PersonId)T
)
```

C Column level constraint

T Table level constraint

1 All constraints need not be provided a name

2 Composite constraint can only be specified as table level constraint

3 Not Null can only be specified as column level constraint

	Clauses
Alter statement 1	`ALTER TABLE Student ADD Address VARCHAR2(20)`
Alter statement 2	`ALTER TABLE Student MODIFY Address VARCHAR2(50)`
Alter statement 3	`ALTER TABLE Student RENAME COLUMN Address TO ResidentialAddress`
Alter statement 4	`ALTER TABLE Student DROP (ResidentialAddress)`
Alter statement 5	`ALTER TABLE Student ADD CONSTRAINT stud_sid_pk PRIMARY KEY (StudentId)`
Alter statement 6	`ALTER TABLE Student DROP CONSTRAINT stud_sid_pk`

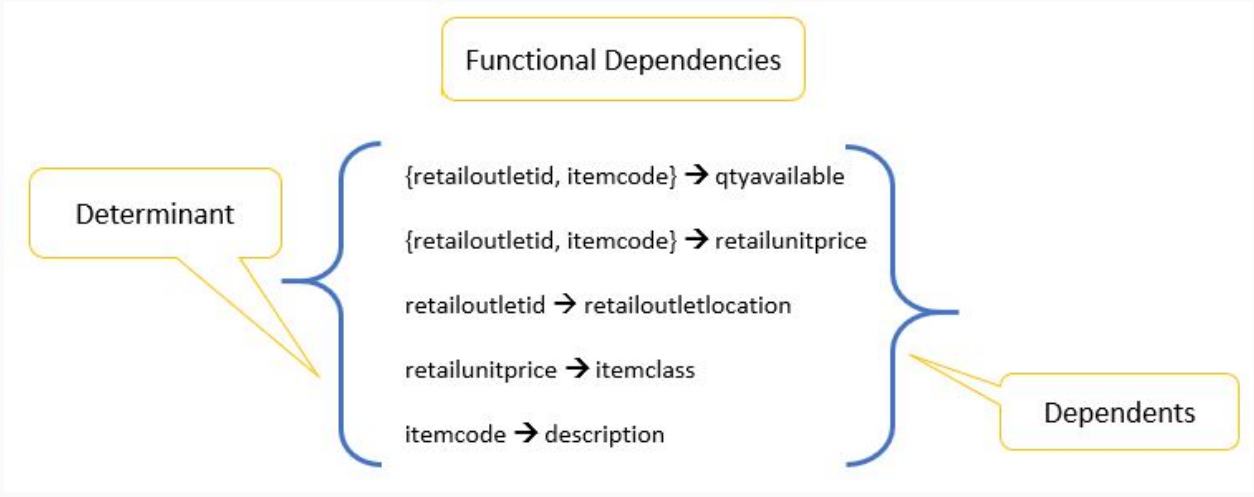
Normalization is the process of reorganizing data in a database to ensure that there is no redundancy of data and data dependencies are logical

FUNCTIONAL DEPENDENCY

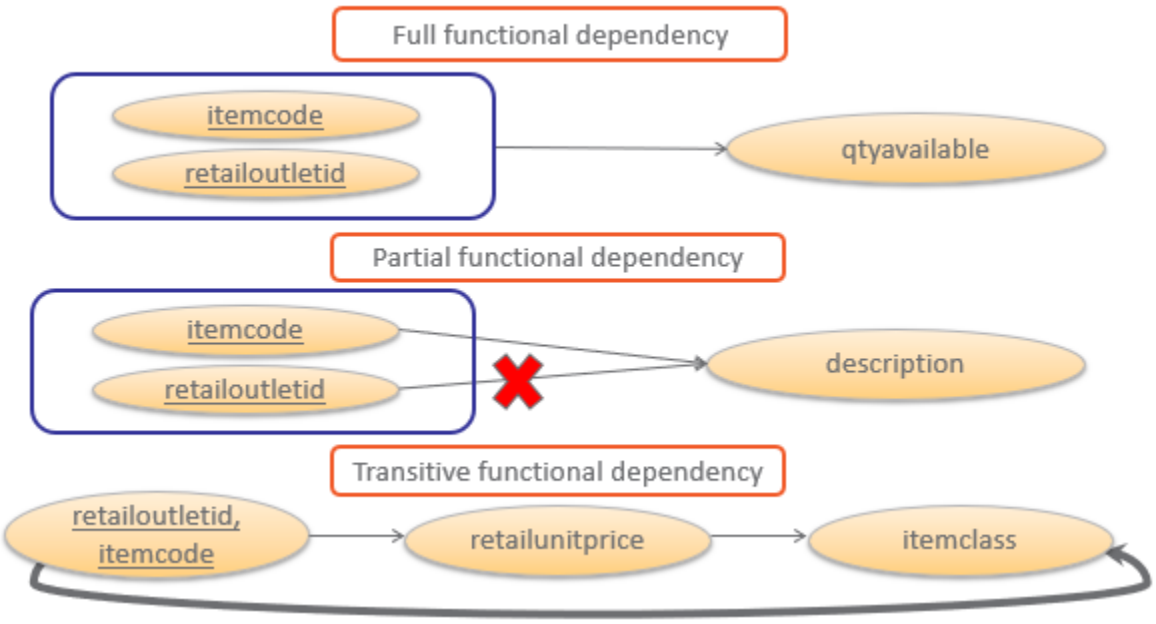
Functional Dependencies in DBMS is a relation between two or more attributes. It can be categorized as -

- Fully Functional Dependency
- Partial Dependency
- Transitive Dependency

The attribute which determines the value of other attributes is known as “Determinant”



- In a relation R, A and B are attributes
- Attribute B is functionally dependent on attribute A if each value of A determines EXACTLY ONE value of B, which is represented as A -> B (A can be composite in nature)
- A is called determinant and B is called dependent



Anomalies which makes the database inconsistent:

1. Update anomaly
2. Insert anomaly
3. Deletion anomaly

Normal Forms" (NF) are the different stages of normalization

- 1 NF (First Normal Form)
- 2 NF (Second Normal Form)
- 3 NF (Third Normal Form)
- BCNF (Boyce -Codd Normal Form)
- 4 NF (Fourth Normal Form)
- 5 NF (Fifth Normal Form)
- 6 NF (Sixth Normal Form)

4NF- 6NF is applicable to multivalued dependencies and complex table scenarios. Based on our course scope, we will discuss 1NF,2NF and 3NF in this module.

What is 1 NF?

A relation R is said to be in 1 NF (First Normal) if and only if:

- All the attributes of R are atomic in nature
- There should not be any multi-valued attribute

What is 2 NF?

A relation R is said to be in 2 NF (Second Normal) form if and only if:

- R is already in 1 NF
- There is no partial dependency in R which exists between non-key attributes and key attributes

What is 3 NF?

A relation R is said to be in 3 NF (Third Normal Form) if and only if:

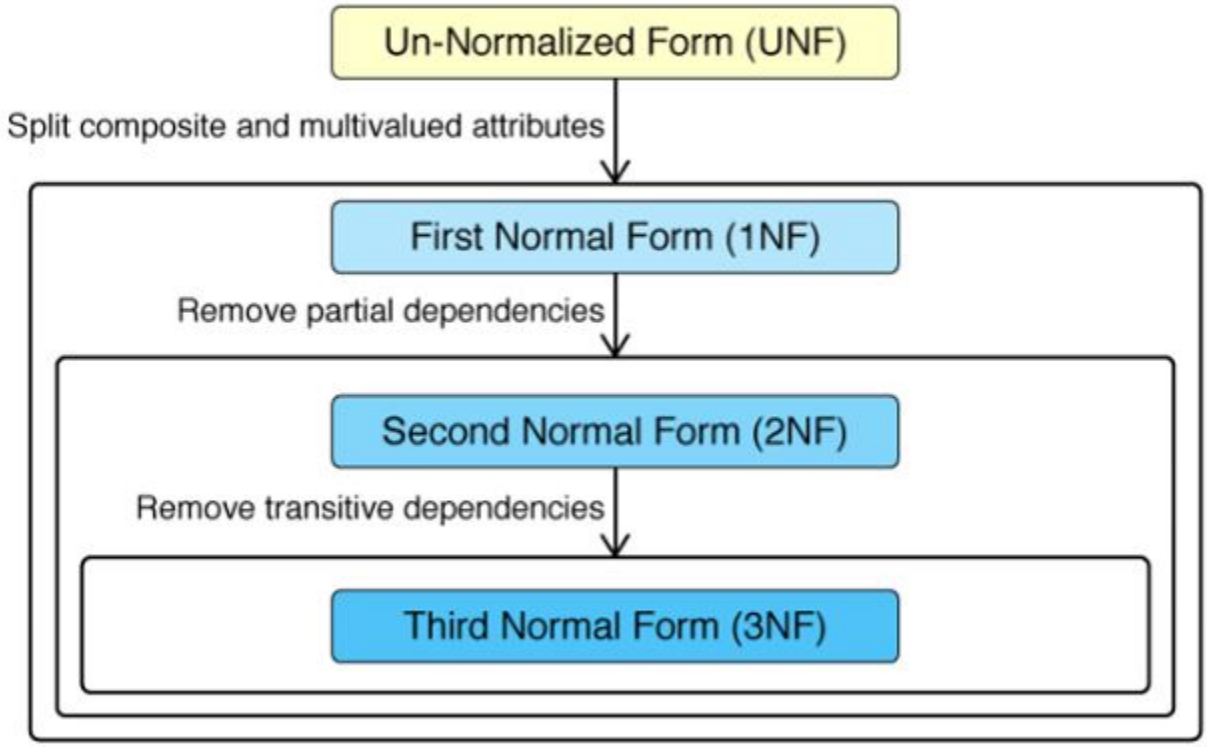
- R is already in 2 NF
- There is no transitive dependency which exists between key attributes and non-key attributes through other non-key attributes

A transitive dependency in a database is an indirect relationship between attributes in the same table that causes a functional dependency.

Guidelines for using normalization:

- Depending on the business requirements, the tables can be normalized up to 2nd normal form or 3rd normal form
- Tables in 3 NF are preferred in applications with extensive data modifications
- Tables in 2 NF are preferred in applications with extensive data retrieval

- Reason: retrieving data from multiple tables is a costly operation
- Converting the tables from higher normal form to lower normal form is called “Denormalization”

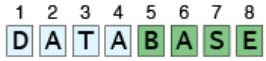


Order of execution of queries:

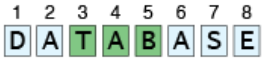
SQL provides many built-in functions in order to accomplish many tasks. Some of the commonly used built-in functions are:

- Numeric functions
- Character functions
- Conversion functions
- Date functions
- Aggregate functions

	Single Row Function	Multi Row Function
Returns	Single Row	Single Row
Operates On	Single Row	Multiple Rows
Used in Clauses	SELECT, WHERE, ORDER BY and HAVING	SELECT, ORDER BY and HAVING clauses



SUBSTR('DATABASE', 5) = 'BASE'



SUBSTR('DATABASE', 3,3) = 'TAB'

All aggregate functions remove ‘null’ case except the Count(*)