Unit III: Database Management

Database Concepts: Introduction to database concepts and its need.

Relational data model: Concept of domain, relation, tuple, attribute, degree, cardinality, key, primary key, candidate key, alternate key and foreign key; Structured Query Language:
General Concepts: Advantages of using SQL, Data Definition Language and Data Manipulation Language;

Data Types: number / decimal, character / varchar , date; SQL commands: CREATE TABLE, DROP TABLE, ALTER TABLE, UPDATE ....SET...., INSERT, DELETE; SELECT, DISTINCT, FROM, WHERE, IN, BETWEEN, LIKE, NULL / IS NULL, ORDER BY,GROUP BY, HAVING;

SQL functions: SUM ( ), AVG ( ), COUNT ( ), MAX ( ) and MIN ( );
Joins: equi-join and natural join

Interface of Python with an SQL database   - Connecting SQL with Python   - Creating Database connectivity Applications   - Performing Insert, Update, Delete queries   - Display data by using fetchone(), fetchall(), rowcount

Database Concepts: Introduction to database concepts and its need.

A database intends to have a collection of data stored together to serve multiple applications as possible. Hence a database is often conceived of as a repository of information needed for running certain functions in a corporation or organization.

Database Management System (DBMS):-Database management system is basically computer based record keeping system.

Advantage of DBMS:-
1. Sharing data
2. Privacy of data
3. Reduce Data Redundancy,
4. Controlled Data Inconsistency,
5. Standardization of data,
6. Data Security.
7. Integrated data.

Database Security:- Data Security refers to protection of data against accidental or intentional discloser to unauthorized persons, or unauthorized modification or destruction.

Database Privacy:-Privacy of data refers to the rights of individuals and organizations to determine for themselves when, how, and to what extent information about them is to be transmitted to others.

DATA ABSTRACTION:- A good database system ensures easy, smooth and efficient data structures in such a way so that every type of database user:
End users (Computer untrained users, at view level)
Application system (at logical level)
Storage system analyst (at internal level or physical implementation)
is able to access desired information efficiently
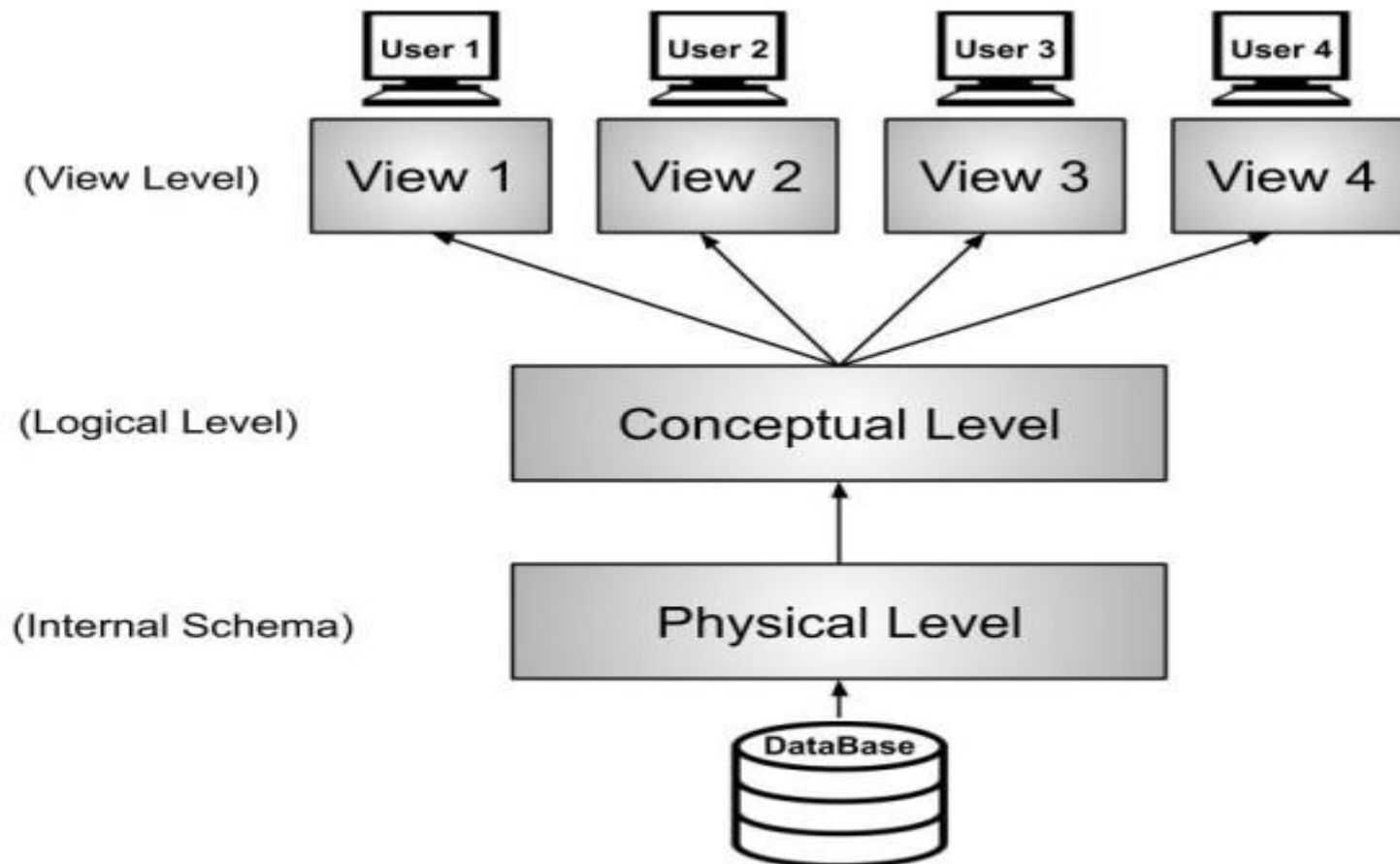
Levels of Data Abstraction

Data Abstraction Type (Various levels of database Implementation)

There are three levels of data abstraction:-

Internal Level (Physical level)
The lowest level of data abstraction, the internal level, is the one closet to physical storage. It describes:-How the data are actually stored in storage device or medium ? What will be the storage technique? What will be the starting address of the database?

Conceptual level
It is next higher level of internal level. In this level data abstraction describe:-What data are actually stored in the data base? What all the constitute the data base? What are the relationships between the data entities?

External level (View level)
This is the level closest to the users and is concerned with the way in which the data are viewed by individual users . It describes the way of viewing information to the concerned user?

Data Independence: The ability to modify a scheme definition in one level without affecting a scheme definition in the next higher level is called Data Independence. There are two
levels of data independence:

***Physical Data Independence:-** It is refers to ability to modify the scheme followed at the physical level without affecting the scheme followed by the conceptual levels.

***Logical Data Independence:-** It is refers to ability to modify the conceptual scheme without causing any changes in the scheme followed at the view levels.
Note:-It is more difficult to achieve logical data independence rather than the physical level data independence.
The abstract data types in modern programming language implement concept of data independence to large extent.

DATA MODELS Model refers to the representation way of data or information in the database management system. The three data models that are used for database management are:-
*   **Relational data model.**
*   **Hierarchical data model.**
*   **Network data model.**

Relational data model:- The relational represents data and relationships among data by a collection of tables known as relations

Relational Model

Activity Code	Activity Name
23	Patching
24	Overlay
25	Crack Sealing

Key = 24

Activity Code	Date	Route No.
24	01/12/01	I-95
24	02/08/01	I-66

Date	Activity Code	Route No.
01/12/01	24	I-95
01/15/01	23	I-495
02/08/01	24	I-66

Network data model:- Network Data model is represented by collection of records and relationships among data are represented by link.
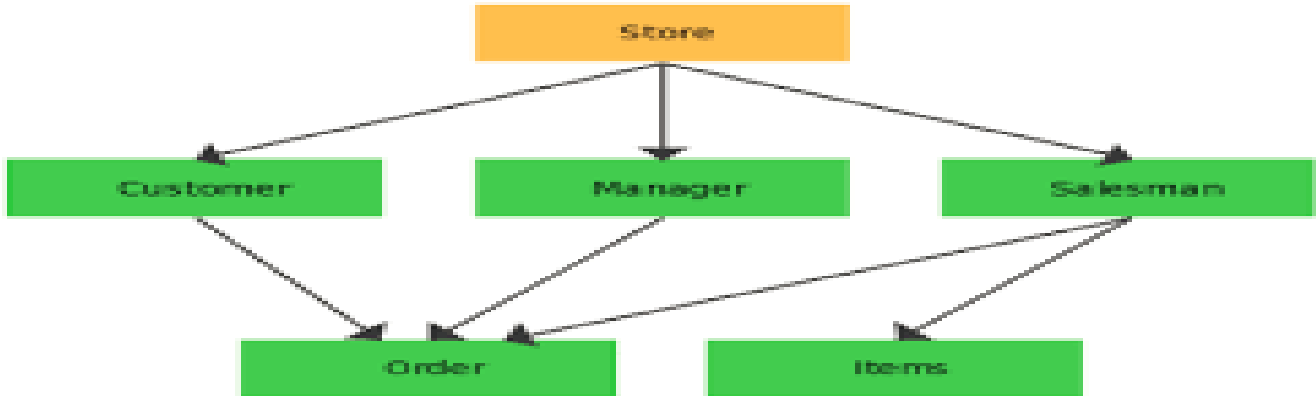
Hierarchical data model:-The hierarchical model is similar to the network model in the sense that data and relationships among data are represented by records and links respectively. It is differs from the network model in that the records are organized as collections of trees rather than arbitrary graphs

.

```
                    ┌──────────────────────┐
                    │        Books         │
                    └──────────┬───────────┘
                    ┌──────────┴───────────┐
                    │     Programming      │
                    └──────────┬───────────┘
          ┌────────────────────┴──────────────────┐
┌─────────┴──────────┐              ┌──────────────┴──────────┐
│     Languages      │              │        Databases        │
└────────────────────┘              └────────────┬────────────┘
                              ┌──────────────────┴─────────────┐
                    ┌─────────┴──────────┐          ┌──────────┴──────────┐
                    │      MongoDB       │          │         dbm         │
                    └────────────────────┘          └─────────────────────┘
```
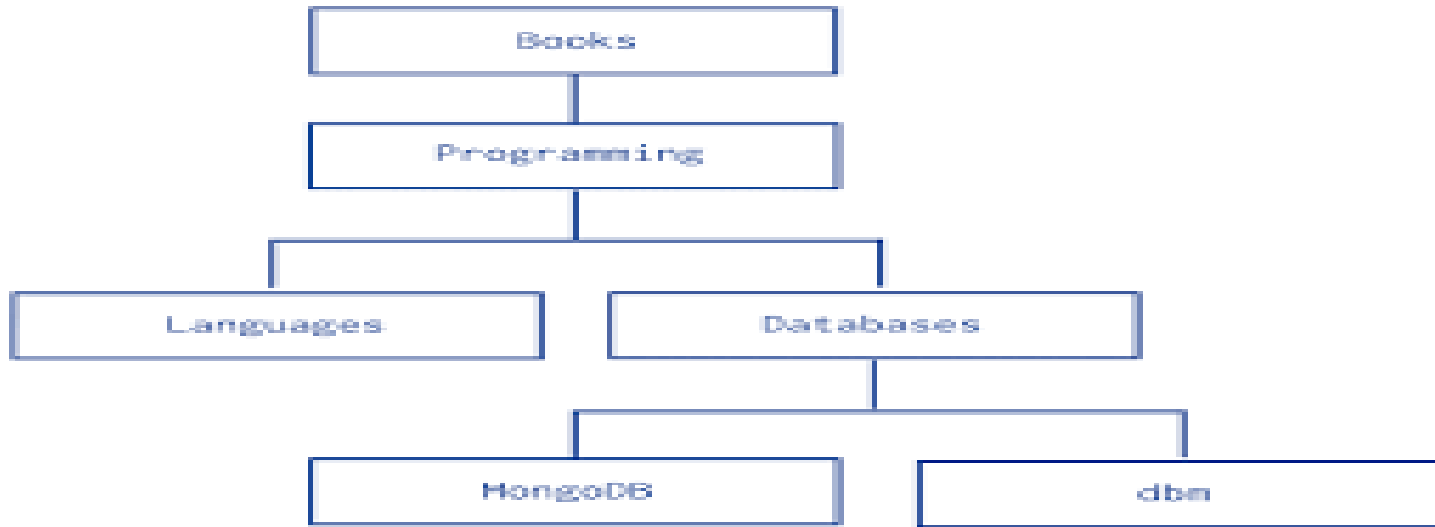
Relational Database Terminology
The relational model was propounded by E.F. Codd of IBM and has since been acknowledged as a very important concept in DBMS technology.
The relational model has established itself as the primary data model for commercial data processing application.
Relation:- A Relation is a table i.e. data arranged in rows and columns.
Domain:- A Domain is pool of values from which the actual values appearing in given column are drawn. Exp. Roll_No# , Student_Name etc.**Tuple:-**The rows of tables (relations) are generally referred to as Tuple.
Attributes:-The columns of tables (relations) are generally referred to as attributes.

Degree:- The number of columns (attributes) in a relation determine the degree of a relation.

Cardinality:- The number of tuples (rows) in a relation is called the cardinality of the relation

Database Keys It is important to be able to specify how rows in a relation are distinguished conceptually, rows are distinct from one another. Database perform following keys:-
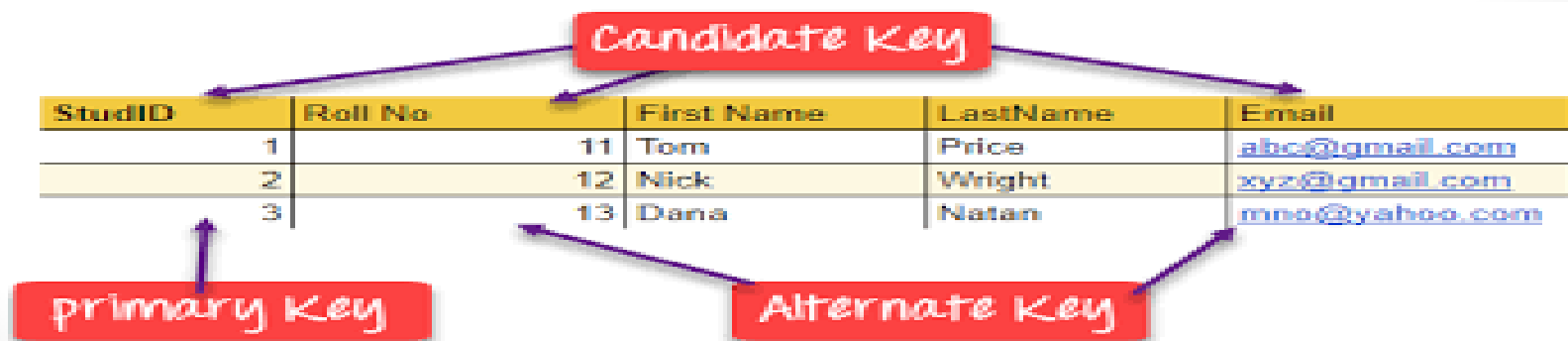
Primary Key Candidate Key Alternate Key Foreign Key Referential Integrity

 Definition of Keys

Primary Key:- A primary key is a set of one or more attributes that can uniquely identify tuples (rows) within the relation/table. For example, Roll_no# is a primary key of "student_rec" table.

Candidate Key:- All attributes combines inside a relation that can serve as a primary key are Candidate Keys as they are candidates for the primary key position.

Alternative Key:- A candidate key that is not the primary key is called an alternative key. For example, student_name is a alternative key in "student_rec" table.

Candidate Key

StudID	Roll No	First Name	LastName	Email
1	11	Tom	Price	abc@gmail.com
2	12	Nick	Wright	xyz@gmail.com
3	13	Dana	Natan	mno@yahoo.com

primary Key

Alternate Key

Foreign Key:- A non key attribute, whose values are derived from the primary key of some other table, is known as Foreign Key in its current table.

EMPLOYEES

Primary Key

"Employee No"
Unique Column Acting as a Foregn Key In "Orders"

SSecurityNo	Employee No	First Name	Last Name	DateOfBirth	Date Employed
AF-23432334	1	Manny	Tomanny	12 Apr 1966	01 May 1999
DQ-65444444	2	Rosanne	Kolumns	21 Mar 1977	01 Jan 2000
GF-54354543	3	Cas	Kade	01 May 1977	01 Apr 2002
JK-34333432	4	Norma	Lyzation	03 Apr 1966	01 Apr 2002
VB-48565444	5	Juan	Tomani	12 Apr 1966	01 Apr 2002
FG-23566553	6	Del	Eats	01 May 1967	01 May 2004

Foreign Key

ORDERS

Primary Key ➡

OrderNo	EmployeeNo	CustomerNo	Supplier	Price	Item
1	1	42	Harrison	$235	Desk
2	4	1	Ford	$234	Chair
3	1	68	Harrison	$415	Table
4	2	112	Ford	$350	Lamp
5	3	42	Ford	$234	Chair
6	2	112	Ford	$350	Lamp
7	2	42	Harrison	$235	Desk

Referential Integrity:-Referential Integrity is a system of rules that a DBMS uses to ensure that relationships between records in related tables are valid, and that users don't accidentally delete or change related data.

Primary Table

CompanyId	CompanyName
1	Apple
2	Samsung

Related Table

CompanyId	ProductId	ProductName
1	1	iPhone
15	2	Mustang

Associated Record ✔

Orphaned Record ✖

MySQL is a freely available open source RDBMS that uses SQL. Information is stored in the form of tables.

Classification of SQL statements:
* **Data definition language(DDL) commands.**
* **Data manipulation language (DML) commands**
* **Transaction Control Language (TCL) commands**
* **Session Control Commands**
* **System Control Commands**

Data definition language(DDL) commands.
DDL allows to add / modify / delete the logical structures which contain the data or which allow users to access / maintain the data (databases, tables, keys, views...). DDL is about "metadata".
Some commands of DDL are:
CREATE – to create table (objects) in the database
ALTER – alters the structure of the database
DROP – delete table from the database
TRUNCATE – remove all records from a table, including all spaces allocated for the records are removed
COMMENT – add comments to the data dictionary
RENAME – rename a table

Data manipulation language (DML) commands

DML allows to add / modify / delete data itself. SELECT is the main DML instruction.

Some commands of DML are:

SELECT – retrieve data from the a database

INSERT – insert data into a table

UPDATE – updates existing data within a table

DELETE – deletes all records from a table, the space for the records remain

****