

## Unit 3: Data Management (DM-1)

---

### Database

**Database:** A Database is an organized collection of facts. In other words we can say that it is a collection of information arranged and presented to serve an assigned purpose. An example of a database is a dictionary where words are arranged alphabetical order. In order to keep database updated we may need to perform operations like adding information, removing information, editing existing information etc. Take a case of phone contacts, in which you save name phone number of friends. This is also a type of database. If you make new friend you save his phone number name it's like adding record to the database.

**Why Database:** Any organization be it Bank, Manufacturing Company, Hospital, University, or Government department requires huge amount of data in some form or other. All such organizations need to collect data, manipulate them and store them for future use. All such type of organization requires data of a number of purpose say:

1. Preparing sales report.
2. Forecasting sales.
3. Preparing accounts reports.
4. Getting medical histories of patients.

Thus we can say that data are very vital corporate resources. The amount of data used these days in organization can be measured in trillions of bytes. The financial investment involved is very high. Many organizations have become critically dependent on the continued and successful operations of database.

**Relational databases:** A Relational database is defined as a database that allows you to group its data items into one or more independent tables that can be related to one another by using fields common to each related table. In Relational Database model, data is organized in the form of rows and columns similar to a table. The tables are referred to as relation in a relational data model. Rows of a table are referred to as tuples and the columns of a table are referred to as attribute. Tables in a relational database is similar to the traditional file system

---

with its records, fields and files. The tables rows contain records and columns contains fields. Each tuple or row has values for the attributes. The attributes describes some property or features of the objects. The full two dimensional table represents a file or “relation”. For example there is a Employee-service relation that describes the entity Employee by the attribute Emp\_code, Name and years of service (years).

EMP_CODE	Name	Years
1234	Harsh	2
1235	Yash	1
1236	Amit	6
1237	Wasim	8

### **Employee-service relation**

Here EMP\_CODE, Name and years are called Attributes. 1234, harsh and 2 are tuples.

**Relation;** All database items and relationships among them are represented in a two dimensional table called a relation. For example Employee-service table shown above is called relation. A database constructed using relations referred to as a relational database. Relational database is constructed from the arrangement of data items. All relations have three components:

1. **Name:** Name is represented by the title or the entity-identifier. For example Employee name and years of services in above table.
2. **Degree:** The number of columns associated with a table or relation among them is called the degree. In Employee-service relation table degree is 3 i.e. there are 3 attributes or fields associated with the table.
3. **Cardinality:** The number of rows in a table is called cardinality. In our example there are 4 records so its cardinality is 4.

**Keys:** The key is the data item that exclusively identifies a record. For example Account\_number, Product\_code, Employee\_number are used as key fields because they specifically identify a record stored in a database.

**Primary Key:** The primary key uniquely identifies each record in a table and must never be

the same for two records. For example Emp-code can be primary key for the entity set Employees. The primary key should be chosen such that its attributes are never very rarely changed. For instance the address field of a person should not be part of the primary key since it's likely to change. EMP\_CODE on the other hand is not changed till he is working in the organization.

**Foreign key:** In a relation the column whose data values correspond to the values of a key column in another relation is called a foreign key. Foreign key of a relation may be the primary key of another relation. For example we have 2 tables "Product" and "Supplier"

