

Data, Database, concepts and Types

- **Data is a collection of facts such as numbers, words, measurement, alphabets, observation and description of things**
- **Database is collection of **related** data in a manner that allows easily access, retrieval, manage, updated and use of that data. A dictionary, a list school children**
- **Database is a collection of related data and data is a collection of facts and figures that can be processed to produce information**
- **Mostly data represents recordable facts.**
- **CONTI.....**

CONTI.....

Database, concepts and Types

- **Data aids in producing information, which is based on facts. For example, if we have data about marks obtained by all students, we can then conclude about toppers and average marks.**
- **A database management system stores data in such a way that it becomes easier to retrieve, manipulate, and produce information.**
- **A data base consists of a collection of tables**
- **Rows** in table are called record (information about a given person, or product or an event)
- **Column** – field (specific piece of information)
- **Information in a database typically stored in table.**
- **Therefore adding, updating, sorting, editing, deleting, etc. are operation we can do in database**

Terminology for data management :

Data : Data describe the entity or object in an enterprise in the form of attributes.

File : A file is a collection of characters describing a useful piece of data or information about the enterprise.

Enterprise : An enterprise is a corporate entity where the requirements occur and applications are developed.

Database : A database is a set of data items in the form of records.

Record : A record is a collection of fields.

Field : A field represents a data item.

Entity : An entity is a collection of attributes of a key object in an enterprise.

Table : A complete set of records constitute a table, organized sequentially.

Operational data : Operational data forms a collection of data items for immediate processing. It is dynamic in nature and accessible by many users frequently.

<i>SNo</i>	<i>title</i>	<i>authors</i>	<i>pub</i>	<i>year</i>	<i>pages</i>	<i>cost</i>	<i>copies</i>
1.	Java book	Norton	PHI	1999	400	200.00	5
2.	C++ primer	Dietel	McGraw	2000	388	300.00	8
3.	DBMS	Korth	M.Hill	1998	660	350.00	7
4.	AI	Robert	Elsevier	1997	352	250.00	4
5.	Data Structures	Kruss	Jain Bros	1996	375	260.00	10
6.	Physics	White	McGraw	2000	292	300.00	10
7.	Chemistry	Shah	AB Pub.	2000	348	320.00	20
8.	Algebra	Joshi	NR Books	2002	441	280.00	15
9.	Logic	Kramer	MIT Press	2000	200	130.00	5
10.	LANs	Mehta	PH Intl.	2001	310	240.00	6

*A dummy set of data items (books in this example)

We can now examine the meaning of the terminology given earlier.

Name of the table : It is books.dbf indicating a database file corresponding to the books in the library.

Attributes : The first row of the table contains eight attributes (no. of columns).

Record : Each row other than the first is a record which is a collection of values of the attributes, called fields.

Format : *SNo*, *year*, *pages*, and *copies* are the attributes where values are in the integer number formats. *title*, *authors*, *pub* are in the character format. *cost* is in real number format.

Field : Each entry such as Korth, 1997, 300.00 etc. are fields.

Row : Each book is being described in a row with eight fields.

Types Database Example

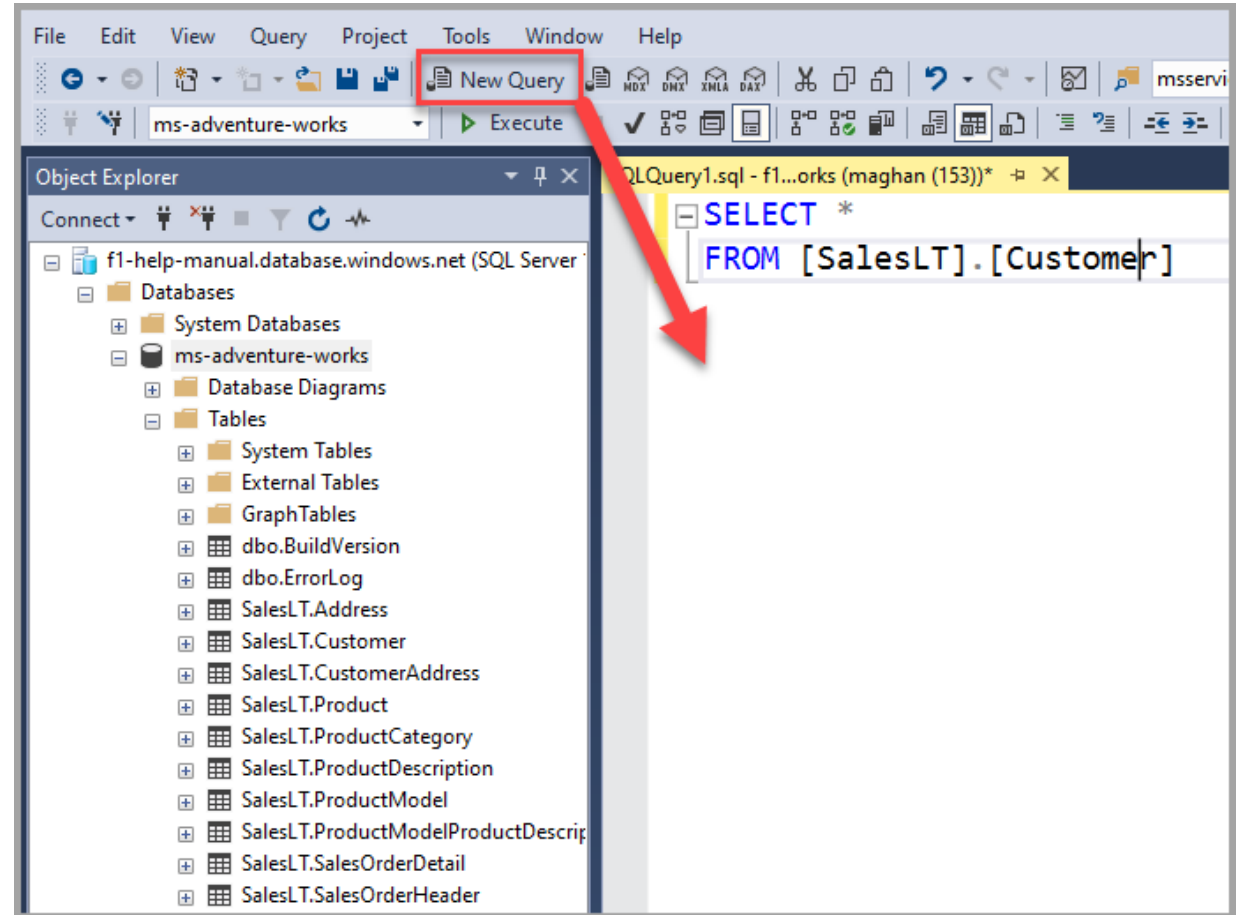
students.dbf

S.NO.	REGISTRATION NO.	NAME	GENDER	DOB
1	200804130001	BODHISATTWA BHAKAT	MALE	xx/xx/xxxx
2	200804130002	GIRIJA SANKAR NAIK	Male	xx/xx/xxxx
3	200804130003	MONALISHA GIRI	FEMALE	xx/xx/xxxx
4	200804130004	NAKKALA VASANTHI	FEMALE	xx/xx/xxxx
5	200804130006	DEBANSHU RANJAN DURGA	MALE	xx/xx/xxxx
6	200804130007	DEBIDAS MANDAL	MALE	xx/xx/xxxx
7	200804130008	DODDA NAMRATHA	FEMALE	xx/xx/xxxx
8	200804130009	IPSITA NANDA	FEMALE	xx/xx/xxxx

- Row 1 = filed Name
- Column 2= Key filed
- Row 1-8 are records
- Reg. no. and DOB is data Items

SQL for Querying the language

- SQL for Querying the Database, SQL is the data manipulating language for querying the database and it is also used for updating the database
- The SQLs of various database system are based on the mathematical formalism called relation algebra where select, sort, join, and meet operations are symbolically dealt for operation



Characteristics of Database

A database is basically a repository of data which is devised in order to support efficient data storage. In addition to data storage, a database also assists in retrieval and maintenance of data stored in it.

Some important characteristics of databases are –

- **Concurrent use** – Many users can access the database concurrently
- **Structured and described data /Self-describing nature of a database** – In addition to data, it also consists of complete definition and description of data
- **Separation of Data and Applications** – Due to this feature it is not required for application software to know about the physical data storage.
- **Data Integrity** – Data remains protected from the unauthorized access and changes

Characteristics of database

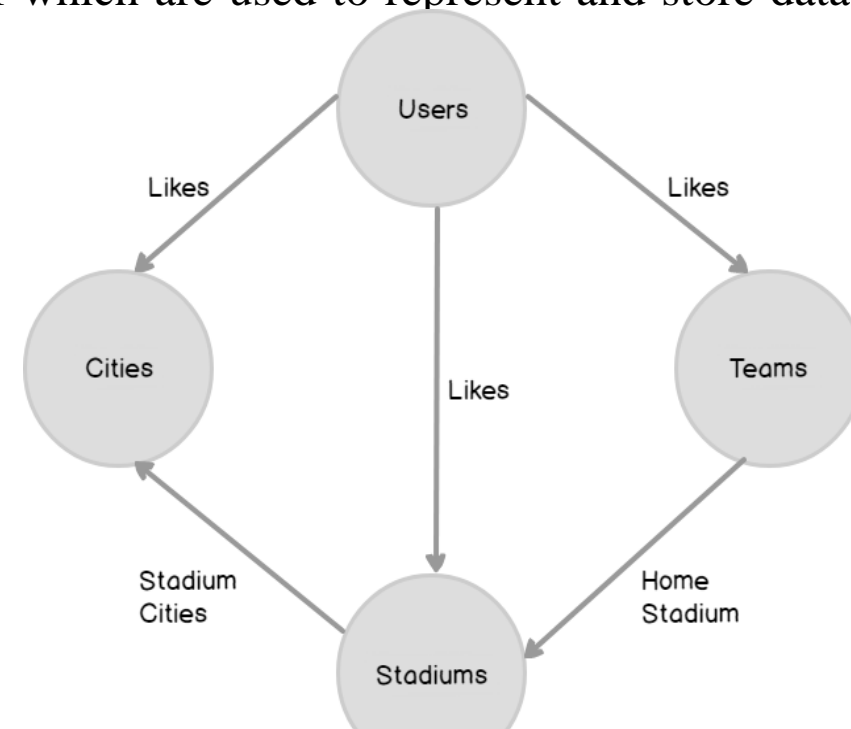
- **Transactions/Transaction processing** – It refers to the features which maintains the atomicity; which means a consistent state.
- **Data Persistence** – It refers to the durability of the data which must not be dependent on any of the system features
- **Database constraint:** a restriction that determines what is allowed to be entered or edited in a table
- **Read-only access:** the ability to read a file but not make changes
- **Metadata:** defines and describes the data and relationships between tables in the database

Types

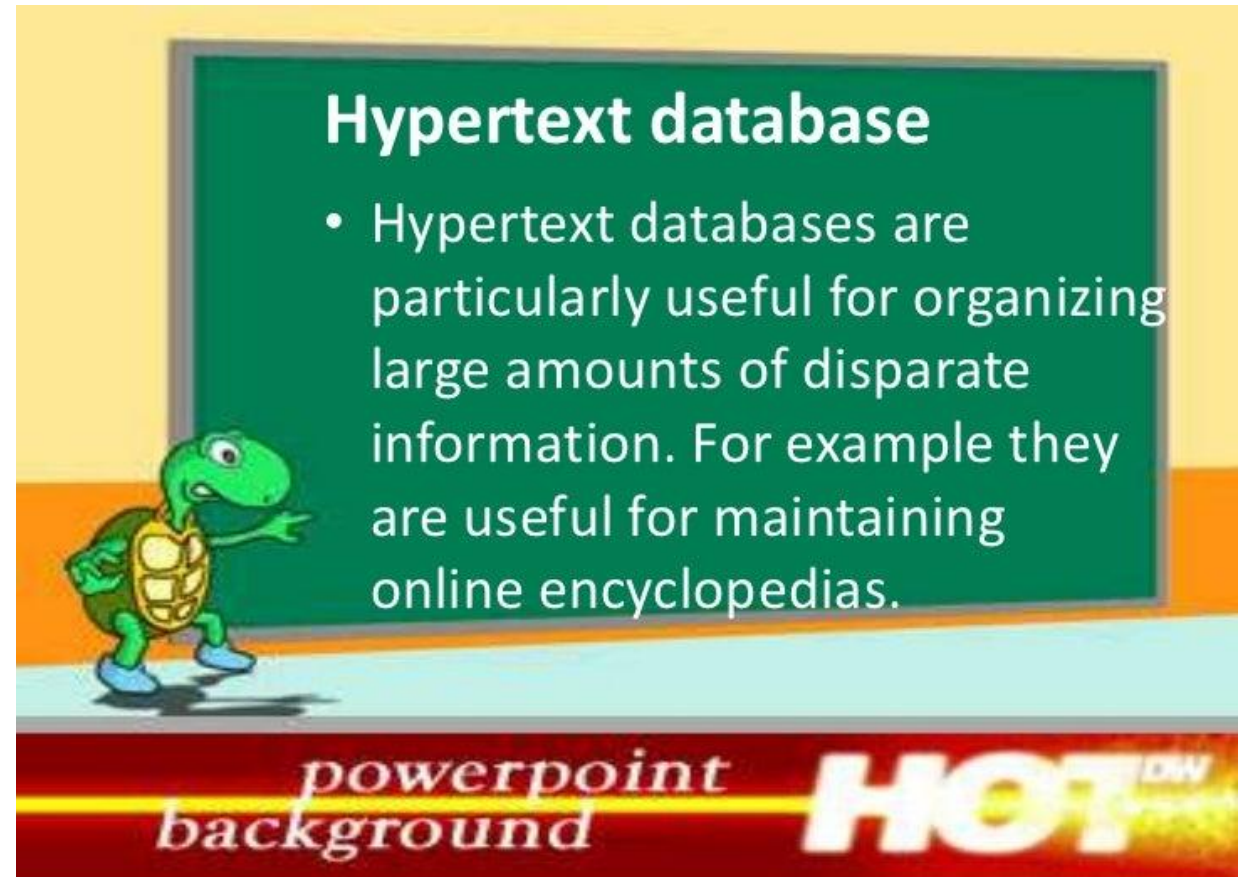
- **Generally, databases can be categorized on the basis of number of users, location, form of data saved in it, etc.**
- **Document Oriented Database** –It does not store data in the form of data table, but in the form of text records. This type of database is suitable for storing dynamic data. **CouchDB and Raven DB** are examples of document databases. It is useful for an application which is document-based. Documents are encoded using some standard formats', all the databases fall in one of these types

- **Embedded Database** – An embedded database runs within an application, and therefore it does not run as a separate application. Unlike general purpose databases, this database is embedded as in line code or linked library.
- It saves time wasted on issues related to installations or maintenance. These types of databases are generally found in the set-top boxes, cell phones, etc. RDM server and RDM Embedded are examples of these types of databases

- **Graph Database** – It is based on the relationship of resources with each other, and no particular resource has any essential importance on the other. These types of graphs help in storing data for a dynamic schema. It provides index free adjunct. In this graph database, each vertex works as a mini-index for its adjacent elements. Info grid type of graph database should be preferred for model flexibility.
- A graph database is defined as **a specialized, single-purpose platform for creating and manipulating graphs**. Graphs contain nodes, edges, and properties, all of which are used to represent and store data in a way that **relational databases are not equipped to do**.



- **Hypertext Database** – These types of databases are used for organizing a large sum of dissimilar information. The type of information is not devised for carrying out numerical analysis. An object is linked with any other object in a hypertext type of database. This kind of database system was invented by Ted Nelson. They are preferred for maintaining online encyclopedias. Unlike traditional databases, it has no regular structure, and therefore the user can reach to the desired information through different ways.



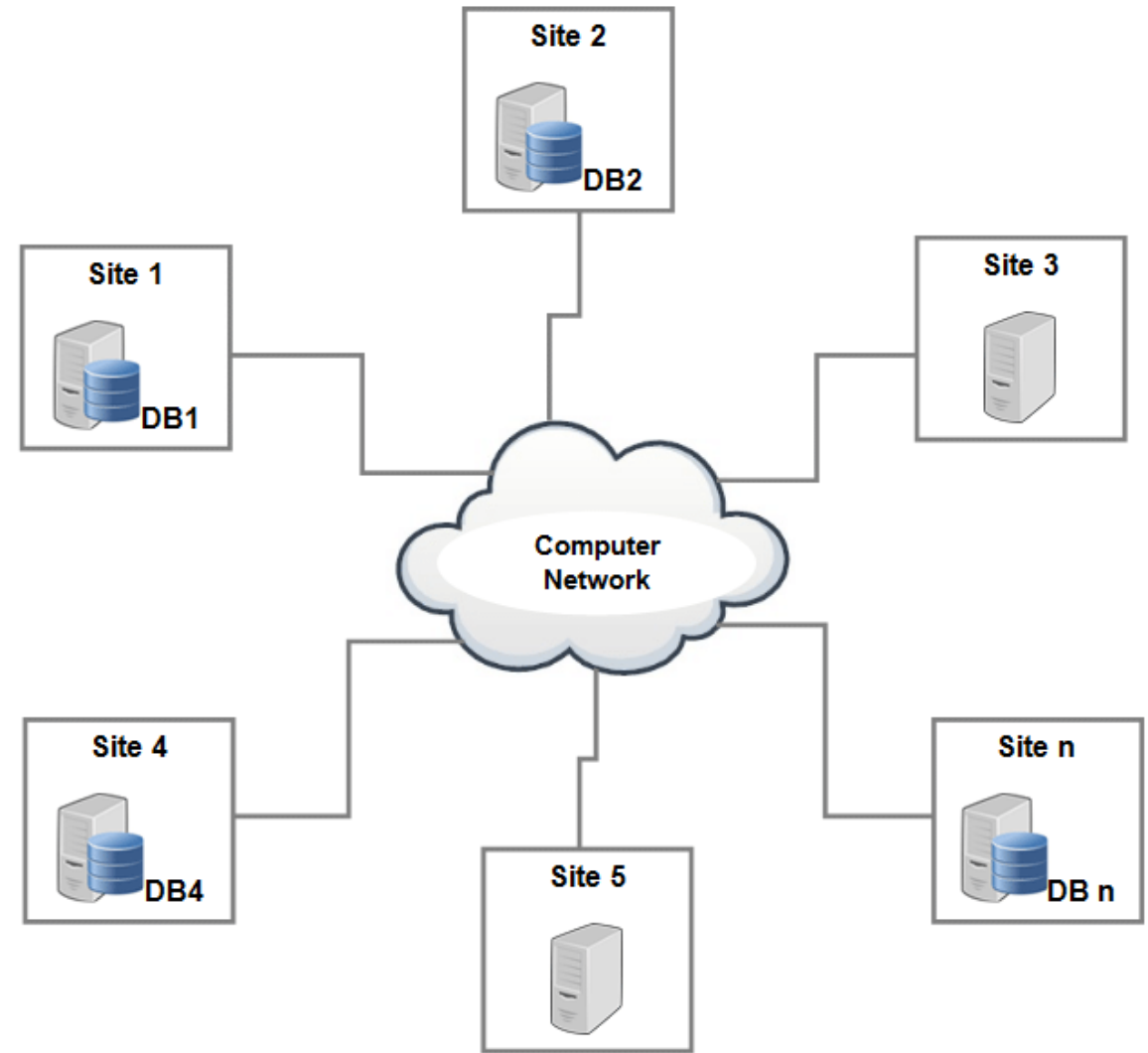
The [icons](#) that you select to view associated objects are called *Hypertext* [links](#) or [buttons](#).

-
- **Operational Database** – It contains data related to the operations going on in an organization or enterprise. Some of the main information it contains are regarding information of employees, data describing transactions, etc. This type of database is updated regularly. It works on the same approach as OLTP. The focus of this database is to record current data. It is often differentiated with the data warehouse.

Operational Database	Data Warehouses
Designed for real time business operations	Designed for analysis of business measures
Designed for common single-line transactions	Are optimized for bulk loads and large complicated queries
Uses data validation tables on a consistent basis	Require no real time data validation
Support thousands of concurrent users.	Support few concurrent users

Source: Inmon W.H (1995)

-
- **Distributed Database** - It consists of a set of databases which are located on different computers, but all these data bases work as one database logically. Therefore, the data can be accessed and modified simultaneously with the help of a network. It is controlled by a local DBMS. It is important to maintain consistency while dealing with this type of arrangement.



- **Flat-File Database** – These are data files in which records hold no structured relationship. Additional information is often required for understanding or interpreting these files. In simple language, if we have one table in a database, it will be referred to as a flat file database. It is useful for storing a small number of records. A spreadsheet application like Excel works as a flat file database.

Flat-File (one table)

Patient Id	Name	D.o.B	Gender	Phone	Doctor Id	Doctor	Room
134	Jeff	4-Jul-1993	Male	7876453	01	Dr Hyde	03
178	David	8-Feb-1987	Male	8635467	02	Dr Jekyll	06
198	Lisa	18-Dec-1979	Female	7498735	01	Dr Hyde	03
210	Frank	29-Apr-1983	Male	7943521	01	Dr Hyde	03
258	Rachel	8-Feb-1987	Female	8367242	02	Dr Jekyll	06