What's a database ?

A database is a collection of data organized in a particular way.
Databases can be of many types such as Flat File Databases, Relational Databases, Distributed Databases etc.

What's an RDBMS ?

This concept was first described around 1970 by Dr. Edgar F. Codd in an IBM research publication called "System R4 Relational".

A relational database uses the concept of linked two-dimensional tables which comprise of rows and columns. A user can draw relationships between multiple tables and present the output as a table again. A user of a relational database need not understand the representation of data in order to retrieve it.

LOGICAL

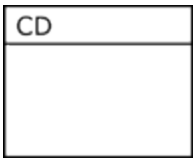
Entity

An **entity** is a thing or object of importance about which data must be captured.

Information about an entity is captured in the form of attributes.

If something is a candidate for being an entity and it has no attributes, it isn't an entity.

Database entities appear in a data model as a box with a title. The title is the name of the entity.

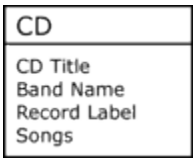


An entity in a data model

An **attribute** describes information about an entity that must be captured, and each attribute describes exactly one entity.

Each entity instance has exactly one value—possibly NULL—for each of its attributes. An attribute value can be a numeric, a character string, a date, a time or some other basic data value.

By common entity naming conventions, an entity name must be singular because each entity names an instance.



The entity with its attributes

E.R.D = Entity Relationship Diagram

PHYSICAL

Table (Entity)

A table consists of rows and columns. Each row in the table represents a collection of related values. Tables are used to hold information about the objects to be represented in the database.

Row (Tuple, Record)

Each row in the table represents a collection of related values of a one object (entity). There is a data of a one customer in one row in Customer table.

A row is a database concept, a tuple is a relational model concept. A record is a little bit outdated term for a tuple or row.

Column (Attribute)

Each column in a table holds a certain kind of data. A Column has a name that describes the data of the column. In the Customer table there are columns e.g. firstname and surname.

A column is a database concept, an attribute is a relational model concept.

Field

A field stores the actual value of an attribute.

Primary key

Primary key is the column (or set of columns) which values uniquely indenfy the row. All primary key fields have a different value in a specific table. A table should have a primary key.

Foreign key

Foreign key is a column whose values refer to the primary key of another table.

What's Database Normalization?

Normalization is the process where a database is designed in a way that removes redundancies, and increases the clarity in organizing data in a database.

Normalization of a database helps in modifying the design at later times and helps in being prepared if a change is required in the database design. Normalization raises the efficiency of the datatabase in terms of management, data storage and scalability.

Now Normalization of a Database is achieved by following a set of rules called 'forms' in creating the database.

1st Normal Form or 1NF:

Each Column Type is Unique. (No repeating Group)

2nd Normal Form or 2NF:

The entity under consideration should already be in the 1NF and all attributes within the entity should depend solely on the entity's unique identifier. (Total Dependent on PK)

3rd Normal Form or 3NF:

The entity should already be in the 2NF and no column entry should be dependent on any other entry (value) other than the key for the table.
If such an entity exists, move it outside into a new table. (No non-key depends on non-key)

Now if these 3NF are achieved, the database is considered normalized.