DATABASE NORMALIZATION

DBA GENESIS

Coming UP

↳ Normalization forms

↳ Normalization example

DBA GENESIS

Normalization forms

- 1st Normal Form

- 2nd Normal Form

- 3rd Normal Form

- Boyce-Codd Normal Form

- 4th Normal Form

- 5th Normal Form

- 6th Normal Form

DBA
GENESIS

There are further enhancements to theory of normalization and it is still being developed. There is even 6th normal form but in any practical scenario, normalization achieves its best shape in 3rd Normal form.

Key terms

- Column – Attribute

- Row – Tuple

- Table – Relation

- Entity – Any real world object that makes sense

DBA GENESIS

Normalization example

Let us look at a library table that maintains all the books they rent out in one single table:

Salutation	Full Name	Address	Books Issued
Mr.	John Mill	710, Street No 4	10X Rule, Titanic, Leader eat last
Mr.	Mani Sharma	12 Lincon Park	Start with why, Billion Dreams, Let's do it
Mr.	Francis Fernando	511, Street No 3	Sell or be sold, Ultimate sales guide

Now let us push this data from various normal forms and see how we can refine the data.

DBA GENESIS

1NF (First Normal Form)

The rules of the first normal form are:

- Each table cell should contain a single/atomic value

- Every record in the table must be unique

Let us first convert the Books_Main_Table into 1NF

DBA
GENESIS

Salutation	Full Name	Address	Books Issued
Mr.	John Mill	710, Street No 4	10X Rule
Mr.	John Mill	710, Street No 4	Titanic
Mr.	John Mill	710, Street No 4	Leader eat last
Mr.	Mani Sharma	12 Lincon Park	Start with why
Mr.	Mani Sharma	12 Lincon Park	Billion Dreams
Mr.	Mani Sharma	12 Lincon Park	Let's do it
Mr.	Francis Fernando	511, Street No 3	Sell or be sold
Mr.	Francis Fernando	511, Street No 3	Ultimate sales guide

As per the 1NF rules, our Books Main Table looks good. Before we proceed with 2NF and 3NF, we need to understand key columns.

DBA
GENESIS

Key / non-key columns

Any column (or group of columns) in a table which can uniquely identify each row is known as key column. For example:

- Phone number

- Email ID

- Student Roll Number

- Employee ID

These are some columns which will always remain unique to every record inside the table.

Such columns are known as key columns inside the table. Any column apart from key columns is known as non-key column.

DBA
GENESIS

Primary key

A primary key is a single column value which uniquely identifies each record in a table. In RDBMS, primary key must satisfy below:

- Primary key must be unique

- Primary key cannot be null

- Every record will have primary key value

DBA
GENESIS

Composite Key

Sometimes its hard to define unique records with one single column. In such cases, we can have two or more columns that uniquely identify each record in a table. Such columns are known as composite key. For example:

- Name + Address

- First Name + DOB + Father Name

Now that we know about key / non-key columns, let us move to 2NF.

2NF (Second Normal Form)

The rules of the second normal form are:

- Table must be in 1NF

- Every non-key attribute must be fully dependent on key attributes

We see that our Books_Main_Table does not have any primary key, in such cases, we will have to introduce a new key column like Membership ID.

To make Books_Main_Table into 2NF, we need to see how columns are closely related:

- Membership ID has a salutation, name and address

- Membership ID has books issued on their name

DBA GENESIS

With this logic in mind, we will have to divide our Books_Main_Table into two table:

Membership ID	Salutation	Full Name	Address
1001	Mr.	John Mill	710, Street No 4
1002	Mr.	Mani Sharma	12 Lincon Park
1003	Mr.	Francis Fernando	511, Street No 3

Membership ID	Books Issued
1001	10X Rule
1001	Titanic
1001	Leader eat last
1002	Start with why
1002	Billion Dreams
1002	Let's do it
1003	Sell or be sold
1003	Ultimate sales guide

If you see the above tables, we have Membership ID in both tables but in Membership_Details_table, it is a primary key column and in Books_Issued_table, it is a non-key column.

Foreign Key

Till now we have seen Primary key and composite key. A foreign key refers to a primary key column of another table. This helps in connecting two tables (and defines a relation between two tables). A foreign key must satisfy below:

- Foreign key column name can be different than primary key column name
- Unlike primary key, then need not be unique (see Books_Issued_Table above)

- Foreign key column can be null even though primary key column cannot

Membership ID	Salutation	Full Name	Address
1001	Mr.	John Mill	710, Street No 4
1002	Mr.	Mani Sharma	12 Lincon Park
1003	Mr.	Francis Fernando	511, Street No 3

Primary Key Column

Foreign Key Column

Membership ID	Books Issued
1001	10X Rule
1001	Titanic
1001	Leader eat last
1002	Start with why
1002	Billion Dreams
1002	Let's do it
1003	Sell or be sold
1003	Ultimate sales guide

DBA GENESIS

Reason for Foreign key

When a user tries to insert a record into Books_Issued_Table and if there is no membership ID exists in Membership_Details_Table, it will be rejected. This way, we maintain data integrity in RDBMS.

If there is no record with Membership ID in the parent table, it will be rejected and database will throw an error.

DBA
GENESIS

3NF (Third Normal Form)
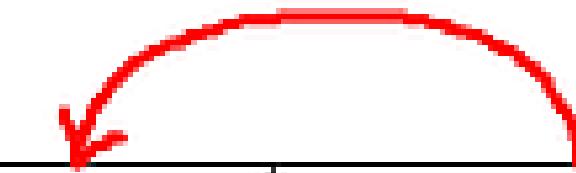
The rules of the third normal form are:

- Data must be in 2NF

- No transitive functional dependencies

What is transitive dependency?

In simple terms, if changing a non-key column causes any other non-key column to change, then its called as transitive dependency.

DBA GENESIS

In our example, if we change Full Name of the customer, it might change Salutation

Membership ID	Salutation	Full Name	Address
1001	Mr.	John Mill	710, Street No 4
1002	Mr.	Mani Sharma	12 Lincon Park
1003	Mr.	Francis Fernando	511, Street No 3

Final 3NF Tables

To move the Membership_Details_Table into 3NF, we need to further divide the table into below:

Membership ID	Salutation ID	Full Name	Address
1001	1	John Mill	710, Street No 4
1002	1	Mani Sharma	12 Lincon Park
1003	1	Francis Fernando	511, Street No 3

Membership ID	Books Issued
1001	10X Rule
1001	Titanic
1001	Leader eat last
1002	Start with why
1002	Billion Dreams
1002	Let's do it
1003	Sell or be sold
1003	Ultimate sales guide

Salutation ID	Salutation
1	Mr.
2	Ms.
3	Mrs.

We have divided the Membership_Details_Table
into a new Salutation_table.

Assignment

If you see the Books_Issued_Table, it still does not have a key column. What do you think should be the key column for the Books_Issued_Table? Or do we need to introduce a new column?

DBA GENESIS

Courses | Projects | Bootcamps