

Ms. Arifa Anjum

Assistant Professor

**St. Aloysius College
(Autonomous) Jabalpur**

What is Normalization in Data Base Management System and why it is required ?

What is Data Base Management System ?

A database management system (DBMS) is a software package designed to define, manipulate, retrieve and manage data in a database. A DBMS generally manipulates the data itself, the data format, field names, record structure and file structure. It also defines rules to validate and manipulate this data.

Definition :-

Normalization is the process of organizing a database.

Why the Normalization is required ?

It is required to reduce the redundancy(repetition) and improve data integrity(consistency).

Normalization is an important part of relational database design, as it helps with the speed, accuracy, and efficiency of the database.

If we never normalize the data so there may be three kind of anomalies can occur.

1. Insert Anomalies

2. Update Anomalies

3. Delete Anomalies

Now, we are considering that this table is not normalized

Eid	Ename	Eaddress	E_Dept
1	Aryan	A1	D001
2	Aryan	A1 A2	D002
3	Mohan	A3	—
<u>4</u>	<u>Rohan</u>	<u>A4</u>	<u>D004</u>

Inconsistency
Update
anomaly

→ If E_Dept can't
accept the null
value so, this
is insert
anomaly

↙

If we the company wants to close the dept no D004, So we will delete the record along with this the information about the emp will also be delete so this is delete anomalies .

Types of Normal Forms

1NF, 2NF, 3NF AND BCNF

What is First Normal Form (1NF) ?

If a relation contain composite or multi-valued attribute, it violates first normal form,

Students

FirstName	LastName	Knowledge
Thomas	Mueller	Java, C++, PHP
Ursula	Meier	PHP, Java
Igor	Mueller	C++, Java

Startsituation

Result after Normalisation



Students

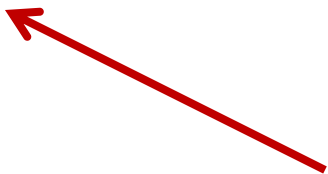
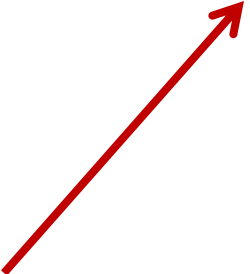
FirstName	LastName	Knowledge
Thomas	Mueller	C++
Thomas	Mueller	PHP
Thomas	Mueller	Java
Ursula	Meier	Java
Ursula	Meier	PHP
Igor	Mueller	Java
Igor	Mueller	C++

Rules for the first normal form :---

- Values of each attribute should be atomic.**
- There should be no composite value.**
- All entries in any column must be of the same kind.**
- Each column must have a unique name.**
- No two rows are identical .**

What is Second Normal Form (2NF) ?

There are the rules for the 2ND Normal form

- 1. The table should be in the first normal form.**
- 2. There should be no partial dependency.**

Functional Dependency

It is a constraint between two sets of attributes in a relation from a Database.

Example : $R(A, B, C, D)$

{

$A \rightarrow B$

$B \rightarrow C$

$C \rightarrow D$

$D \rightarrow A$

}

Determinant

Dependent

Functional dependency

StudentID → Student Name



Student ID	Student Name	Subject ID	Subject result
1	Bill	IT10	90%
2	Mike	IT6	60%
2	Mike	BI4	90%
3	Megan	LC2	75%
4	Bill	BI8	65%

SECOND NORMAL FORM

Id	Name	Subjects
1289	Ramesh Sawant	Math, Science
5678	Shruti Shah	English



Id	Name	Subjects
1289	Ramesh Sawant	Math
1289	Ramesh Sawant	Science
5678	Shruti Shah	English



Id	Subjects
1289	Math
1289	Science
5678	English

Id	Name
1289	Ramesh Sawant
1289	Ramesh Sawant
5678	Shruti Shah

What is Partial Dependency?

Partial Dependency occurs when a non-prime attribute is functionally dependent on part of a candidate key.

The 2nd Normal Form (2NF) eliminates the Partial Dependency

StudentID	ProjectNo	StudentName	ProjectName
S01	199	Katie	Geo Location
S02	120	Ollie	Cluster Exploration

StudentID = Unique ID of the student

StudentName = Name of the student

ProjectNo = Unique ID of the project

ProjectName = Name of the project

To remove Partial Dependency and violation on 2NF, decompose the tables:
<StudentInfo>

StudentID	ProjectNo	StudentName
S01	199	Katie
S02	120	Ollie

<ProjectInfo>

ProjectNo	ProjectName
199	Geo Location
120	Cluster Exploration

Third Normal Form (3NF) ?

A database is in third normal form if it satisfies the following conditions:

- It is in second normal form
- There is no transitive functional dependency

Student Table

Stu_id	Name	Reg_no	Branch	Address

Score Table

Score_id	Student_id	Subject_id	Marks	Teacher
	Composite Primary key			

Non prime attribute depends on the part of the candidate key , so we will break this table into subject table.

Subject Table

Subject_id	Subject_name	Teacher

If we add the exam_name & total column in score table.

Score Table

Score_id	Student_id	Subject_id	Marks	Teacher	Exam_name	Total Marks

Here , the Exam_name attribute will depend on Student_id & Subject_id but the Total_Marks attribute will depend on Exam_name which is not a Primary key .This is the transitive dependency.

So we will break the table with name Exam table. .

Exam Table

Exam_Name	Total_marks
-----------	-------------

What is BCNF?

A relation **R** is said to be in BCNF ,if and only if every determinant is a primary key.

Now, considering an example.

Id	E_Name	Qualification	Grade	Did	Dname
1	Satish	B.E.	C	20	EC
2	Savita	M.E.	B	30	CE
3	Mahesh	P.hd.	A	10	IT

1. **Id , Qualification -----> Grade**
2. **Id -----> E_name,Did**
3. **Did -----> Dname**

Here , we will make these three tables

id	Qualification	Grade
1	BE	C
2	ME	B

Id	Ename	Did
1	Satish	20
2	Savita	30

Did	Dname
10	EC
20	ME

Thank you