

[Suggest a Topic](#)[Login](#)[Write an Article](#)

## Database Normalization | Normal Forms



**Prerequisite** – Database normalization and functional dependency concept.

**Normalization** is the process of minimizing **redundancy** from a relation or set of relations. Redundancy in relation may cause insertion, deletion and updation anomalies. So, it helps to minimize the redundancy in relations. **Normal forms** are used to eliminate or reduce redundancy in database tables.

### 1. First Normal Form –

If a relation contain composite or multi-valued attribute, it violates first normal form or a relation is in first normal form if it does not contain any composite or multi-valued attribute. A relation is in first normal form if every attribute in that relation is **singled valued attribute**.

- **Example 1** – Relation STUDENT in table 1 is not in 1NF because of multi-valued attribute STUD\_PHONE. Its decomposition into 1NF has been shown in table 2.



STUD_NO	STUD_NAME	STUD_PHONE	STUD_STATE	STUD_COUNTRY
1	RAM	9716271721, 9871717178	HARYANA	INDIA
2	RAM	9898297281	PUNJAB	INDIA
3	SURESH		PUNJAB	INDIA

Table 1

Conversion to first normal form

STUD_NO	STUD_NAME	STUD_PHONE	STUD_STATE	STUD_COUNTRY
1	RAM	9716271721	HARYANA	
1	RAM	9871717178	HARYANA	INDIA
2	RAM	9898297281	PUNJAB	INDIA
3	SURESH		PUNJAB	INDIA

Table 2

### ■ Example 2 –

ID	Name	Courses
1	A	c1, c2
2	E	c3
3	M	C2, c3

In the above table Course is a multi valued attribute so it is not in 1NF.

Below Table is in 1NF as there is no multi valued attribute

ID	Name	Course
1	A	c1
1	A	c2
2	E	c3
3	M	c1
3	M	c2

## 2. Second Normal Form –

To be in second normal form, a relation must be in first normal form and relation must not contain any partial dependency. A relation is in 2NF iff it has **No Partial Dependency**, i.e., no non-prime attribute (attributes which are not part of any candidate key) is dependent on any proper subset of any candidate key of the table.



STUD_NO	COURSE_NO	COURSE_NAME
1	C1	DBMS
2	C2	Computers Network
1	C2	Computers Network

**Table 3**

**Partial Dependency** – If proper subset of candidate key determines non-prime attribute, it is called partial dependency.

- **Example 1** – In relation STUDENT\_COURSE given in Table 3,

FD set: {COURSE\_NO → COURSE\_NAME}  
 Candidate Key: {STUD\_NO, COURSE\_NO}

In FD COURSE\_NO → COURSE\_NAME, COURSE\_NO (proper subset of candidate key) is determining COURSE\_NAME (non-prime attribute). Hence, it is partial dependency and relation is not in second normal form.

To convert it to second normal form, we will decompose the relation STUDENT\_COURSE (STUD\_NO, COURSE\_NO, COURSE\_NAME) as :

STUDENT\_COURSE (STUD\_NO, COURSE\_NO)  
 COURSE (COURSE\_NO, COURSE\_NAME)

Note – This decomposition will be lossless join decomposition as well as dependency preserving.

- **Example 2** – Consider following functional dependencies in relation R (A, B, C, D)

AB → C [A and B together determine C]  
 BC → D [B and C together determine D]

In the above relation, AB is the only candidate key and there is no partial dependency, i.e., any proper subset of AB doesn't determine any non-prime attribute.

### 3. Third Normal Form –

A relation is in third normal form, if there is **no transitive dependency** for non-prime attributes is it is in second normal form.

A relation is in 3NF iff **at least one of the following condition holds** in every non-trivial function dependency  $X \rightarrow Y$

1. X is a super key.
2. Y is a prime attribute (each element of Y is part of some candidate key).



STUD_NO	STUD_NAME	STUD_STATE	STUD_COUNTRY	STUD_AGE
1	RAM	HARYANA	INDIA	20
2	RAM	PUNJAB	INDIA	19
3	SURESH	PUNJAB	INDIA	21

Table 4

**Transitive dependency** – If  $A \rightarrow B$  and  $B \rightarrow C$  are two FDs then  $A \rightarrow C$  is called transitive dependency.

- **Example 1** – In relation STUDENT given in Table 4,

FD set: {STUD\_NO  $\rightarrow$  STUD\_NAME, STUD\_NO  $\rightarrow$  STUD\_STATE, STUD\_NO  $\rightarrow$  STUD\_COUNTRY, STUD\_NO  $\rightarrow$  STUD\_AGE, STUD\_STATE  $\rightarrow$  STUD\_COUNTRY}

Candidate Key: {STUD\_NO}

For this relation in table 4, STUD\_NO  $\rightarrow$  STUD\_STATE and STUD\_STATE  $\rightarrow$  STUD\_COUNTRY are true.

So STUD\_COUNTRY is transitively dependent on STUD\_NO. It violates third normal form. To convert it in third normal form, we will decompose the relation STUDENT (STUD\_NO, STUD\_NAME, STUD\_PHONE, STUD\_STATE, STUD\_COUNTRY, STUD\_AGE) as:

STUDENT (STUD\_NO, STUD\_NAME, STUD\_PHONE, STUD\_STATE, STUD\_AGE)

STATE\_COUNTRY (STATE, COUNTRY)

- **Example 2** – Consider relation R(A, B, C, D, E)

A  $\rightarrow$  BC,

CD  $\rightarrow$  E,

B  $\rightarrow$  D,

E  $\rightarrow$  A

All possible candidate keys in above relation are {A, E, CD, BC} All attribute are on right sides of all functional dependencies are prime.

#### 4. Boyce-Codd Normal Form (BCNF) –

A relation R is in BCNF if R is in Third Normal Form and for every FD, LHS is super key. A relation is in BCNF iff in every non-trivial functional dependency  $X \rightarrow Y$ , X is a super key.

- **Example 1** – Find the highest normal form of a relation R(A,B,C,D,E) with FD set as {BC $\rightarrow$ D, AC $\rightarrow$ BE, B $\rightarrow$ E}

Step 1. As we can see, (AC)<sup>+</sup> = {A,C,B,E,D} but none of its subset can determine all attribute of relation, So AC will be candidate key. A or C can't be derived from any other attribute of the relation, so there will be only 1 candidate key {AC}.

Step 2. Prime attribute are those attribute which are part of candidate key {A,C} in this example and others will be non-prime {B,D,E} in this example.

Step 3. The relation R is in 1st normal form as a relational DBMS does not allow multi-valued or composite attribute.

The relation is in 2nd normal form because BC $\rightarrow$ D is in 2nd normal form (BC is not proper subset of

candidate key AC) and  $AC \rightarrow BE$  is in 2nd normal form (AC is candidate key) and  $B \rightarrow E$  is in 2nd normal form (B is not a proper subset of candidate key AC).

The relation is not in 3rd normal form because in  $BC \rightarrow D$  (neither BC is a super key nor D is a prime attribute) and in  $B \rightarrow E$  (neither B is a super key nor E is a prime attribute) but to satisfy 3rd normal form, either LHS of an FD should be super key or RHS should be prime attribute.

So the highest normal form of relation will be 2nd Normal form.

▪ **Example 2** –For example consider relation R(A, B, C)

$A \rightarrow BC$ ,

$B \rightarrow$

A and B both are super keys so above relation is in BCNF.

**Key Points –**

1. BCNF is free from redundancy.
2. If a relation is in BCNF, then 3NF is also satisfied.
3. If all attributes of relation are prime attribute, then the relation is always in 3NF.
4. A relation in a Relational Database is always and at least in 1NF form.
5. Every Binary Relation ( a Relation with only 2 attributes ) is always in BCNF.
6. If a Relation has only singleton candidate keys( i.e. every candidate key consists of only 1 attribute), then the Relation is always in 2NF( because no Partial functional dependency possible).
7. Sometimes going for BCNF form may not preserve functional dependency. In that case go for BCNF only if the lost FD(s) is not required, else normalize till 3NF only.
8. There are many more Normal forms that exist after BCNF, like 4NF and more. But in real world database systems it's generally not required to go beyond BCNF.

**Exercise 1:** Find the highest normal form in R (A, B, C, D, E) under following functional dependencies.

$ABC \twoheadrightarrow D$   
 $CD \twoheadrightarrow AE$

**Important Points** for solving above type of question.

- 1) It is always a good idea to start checking from BCNF, then 3 NF and so on.
- 2) If any functional dependency satisfied a normal form then there is no need to check for lower normal form. For example,  $ABC \rightarrow D$  is in BCNF (Note that ABC is a super key), so no need to check this dependency for lower normal forms.

Candidate keys in given relation are {ABC, BCD}

**BCNF:**  $ABC \rightarrow D$  is in BCNF. Let us check  $CD \rightarrow AE$ , CD is not a super key so this dependency is not in BCNF. So, R is not in BCNF.

**3NF:**  $ABC \rightarrow D$  we don't need to check for this dependency as it already satisfied BCNF. Let us consider  $CD \rightarrow AE$ . Since E is not a prime attribute, so relation is not in 3NF.



**2NF:** In 2NF, we need to check for partial dependency. CD which is a proper subset of a candidate key and it determine E, which is non prime attribute. So, given relation is also not in 2 NF. So, the highest normal form is 1 NF.

### GATE CS Corner Questions

Practicing the following questions will help you test your knowledge. All questions have been asked in GATE in previous years or in GATE Mock Tests. It is highly recommended that you practice them.

1. [GATE CS 2012, Question 2](#)
2. [GATE CS 2013, Question 54](#)
3. [GATE CS 2013, Question 55](#)
4. [GATE CS 2005, Question 29](#)
5. [GATE CS 2002, Question 23](#)
6. [GATE CS 2002, Question 50](#)
7. [GATE CS 2001, Question 48](#)
8. [GATE CS 1999, Question 32](#)
9. [GATE IT 2005, Question 22](#)
10. [GATE IT 2008, Question 60](#)
11. [GATE CS 2016 \(Set 1\), Question 31](#)

See [Quiz on Database Normal Forms](#) for all previous year questions.

This article is contributed by **Sonal Tuteja**. If you like GeeksforGeeks and would like to contribute, you can also write an article using [contribute.geeksforgeeks.org](https://www.geeksforgeeks.org/contribute) or mail your article to [contribute@geeksforgeeks.org](mailto:contribute@geeksforgeeks.org). See your article appearing on the GeeksforGeeks main page and help other Geeks.

Please write comments if you find anything incorrect, or you want to share more information about the topic discussed above.

Practice Tags :

DBMS

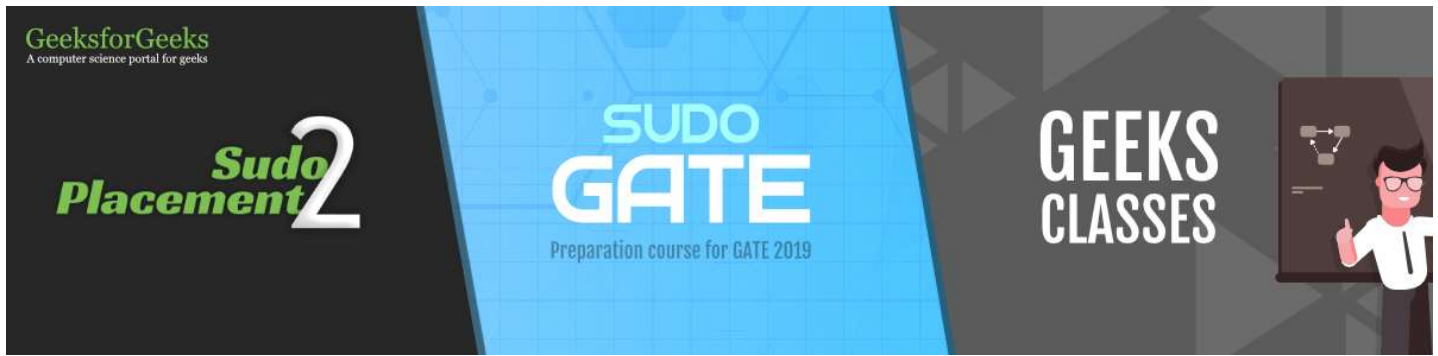
Article Tags :

DBMS

GATE CS

DBMS-Normalization





Please write to us at [contribute@geeksforgeeks.org](mailto:contribute@geeksforgeeks.org) to report any issue with the above content.

[Login to Improve this Article](#)

## Recommended Posts:

[Database Management System | Dependency Preserving Decomposition](#)

[Database Normalization | Introduction](#)

[ACID Properties in DBMS](#)

[DBMS | Keys in Relational Model \(Candidate, Super, Primary, Alternate and Foreign\)](#)

[Functional Dependency and Attribute Closure](#)

[CTE in SQL](#)

[Paytm Interview experience for FTE \(On-Campus\)](#)

[DBMS | Join operation Vs nested query](#)

[Amadeus Labs R & D | On Campus \(freshers\) | Full time+Internship](#)

[DBMS | Types of OLAP Systems](#)

(Login to Rate)

**3.5** Average Difficulty : **3.5/5.0**  
Based on **35** vote(s)

☐ Add to TODO List

☐ Mark as DONE

[Feedback](#)

[Add Notes](#)

[Improve Article](#)

Basic

Easy

Medium

Hard

Expert

Writing code in comment? Please use [ide.geeksforgeeks.org](https://ide.geeksforgeeks.org), generate link and share the link here.

[Load Comments](#)

[Share this post!](#)



## A computer science portal for geeks

710-B, Advant Navis Business Park,  
Sector-142, Noida, Uttar Pradesh - 201305  
[feedback@geeksforgeeks.org](mailto:feedback@geeksforgeeks.org)

### COMPANY

About Us  
Careers  
Privacy Policy  
Contact Us

### PRACTICE

Company-wise  
Topic-wise  
Contests  
Subjective Questions

### LEARN

Algorithms  
Data Structures  
Languages  
CS Subjects  
Video Tutorials

### CONTRIBUTE

Write an Article  
Write Interview Experience  
Internships  
Videos

@geeksforgeeks, Some rights reserved

