

# Database Normalization

If you've been working with databases for a while, chances are you've heard the term normalization. Perhaps someone's asked you "Is that database normalized?" or "Is that in BCNF?" All too often, the reply is "Yeah." Normalization is often brushed aside as a luxury that only academics have time for. However, knowing the principles of normalization and applying them to your daily database design tasks really isn't all that complicated and it could drastically improve the performance of your DBMS.

In this article, we'll introduce the concept of normalization and take a brief look at the most common normal forms. Future articles will provide in-depth explorations of the normalization process.

So, what is normalization? Basically, it's the process of efficiently organizing data in a database.

There are two goals of the normalization process: eliminate redundant data (for example, storing the same data in more than one table) and ensure data dependencies make sense (only storing related data in a table). Both of these are worthy goals as they reduce the amount of space a database consumes and ensure that data is logically stored.

The database community has developed a series of guidelines for ensuring that databases are normalized. These are referred to as normal forms and are numbered from one (the lowest form of normalization, referred to as first normal form or 1NF) through five (fifth normal form or 5NF). In practical applications, you'll often see 1NF, 2NF, and 3NF along with the occasional 4NF. Fifth normal form is very rarely seen and won't be discussed in this article.

Before we begin our discussion of the normal forms, it's important to point out that they are guidelines and guidelines only. Occasionally, it becomes necessary to stray from them to meet practical business requirements. However, when variations take place, it's extremely important to evaluate any possible ramifications they could have on your system and account for possible inconsistencies. That said, let's explore the normal forms.

First normal form (1NF) sets the very basic rules for an organized database:

- Eliminate duplicative columns from the same table.
- Create separate tables for each group of related data and identify each row with a unique column or set of columns (the primary key).

Second normal form (2NF) further addresses the concept of removing duplicative data:

- Remove subsets of data that apply to multiple rows of a table and place them in separate tables.
- Create relationships between these new tables and their predecessors through the use of foreign keys.

Third normal form (3NF) goes one large step further:

- Remove columns that are not dependent upon the primary key.

Finally, fourth normal form (4NF) has one requirement:

- A relation is in 4NF if it has no multi-valued dependencies.

Remember, these normalization guidelines are cumulative. For a database to be in 2NF, it must first fulfill all the criteria of a 1NF database.

**Source:** <http://www.go4expert.com/articles/database-normalization-t299/>