About.com Databases

Database Normalization Basics

Normalizing your Database

By Mike Chapple

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[1] normalized?" or "Is that in BCNF[2]?" All too often, the reply is "Uh, 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.

What is Normalization?

Normalization is the process of efficiently organizing data in a database. There are two goals of the normalization process: eliminating redundant data (for example, storing the same data in more than one table[3]) and ensuring data dependencies[4] 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 Normal Forms

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[5] or 1NF) through five (fifth normal form or 5NF). In practical applications, you'll often see 1NF[6], 2NF[7], and 3NF[8] 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)

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

- Eliminate duplicative columns[9] from the same table.
- Create separate tables for each group of related data and identify each row[10] with a unique column or set of columns (the primary key[11]).

For more details, read **Putting your Database in First Normal Form[12]**

Second Normal Form (2NF)

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

- Meet all the requirements of the first normal form.
- 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[13].

For more details, read **Putting your Database in Second Normal Form[14]**

Third Normal Form (3NF)

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

- Meet all the requirements of the second normal form.
- Remove columns that are not dependent[15] upon the primary key.

For more details, read **Putting your Database in Third Normal Form[16]**

Boyce-Codd Normal Form (BCNF or 3.5NF)

The Boyce-Codd Normal Form, also referred to as the "third and half (3.5) normal form", adds one more requirement:

- Meet all the requirements of the third normal form.
- Every determinant must be a candidate key.[17]

For more details, read [18]Putting your Database in Boyce Codd Normal Form[19]

Fourth Normal Form (4NF)

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

- Meet all the requirements of the third normal form.
- A relation is in 4NF if it has no multi-valued dependencies[20].

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

Should I Normalize?

While database normalization is often a good idea, it's not an absolute requirement. In fact, there are some cases where deliberately violating the rules of normalization is a good practice. For more on this topic, read Should I Normalize My Database?[21].

If you'd like to ensure your database is normalized, explore our other articles in this series:

- **Database Normalization Basics**[22]
- **Putting your Database in First Normal Form**[23]
- **Putting your Database in Second Normal Form**[24]
- **Putting your Database in Third Normal Form**[25]

If you want to receive notifications of new database articles posted on this site, Subscribe to our newsletter[26]

This About.com page has been optimized for print. To view this page in its original form, please visit: http://databases.about.com/od/specificproducts/a/normalization.htm

Links in this article:
1. http://databases.about.com/library/glossary/bldef-database.htm
2. http://databases.about.com/library/glossary/bldef-bcnf.htm
3. http://databases.about.com/library/glossary/bldef-table.htm
4. http://databases.about.com/od/specificproducts/a/Database-Dependency.htm
5. http://databases.about.com/od/specificproducts/l/aa1nf.htm
6. http://databases.about.com/od/specificproducts/l/aa1nf.htm
7. http://databases.about.com/od/specificproducts/a/2nf.htm
8. http://databases.about.com/od/specificproducts/l/aa3nf.htm
9. http://databases.about.com/library/glossary/bldef-column.htm
10. http://databases.about.com/library/glossary/bldef-row.htm
11. http://databases.about.com/library/glossary/bldef-primarykey.htm
12. http://databases.about.com/od/specificproducts/l/aa1nf.htm
13. http://databases.about.com/library/glossary/bldef-foreignkey.htm
14. http://databases.about.com/od/specificproducts/a/2nf.htm
15. http://databases.about.com/od/specificproducts/a/Database-Dependency.htm
16. http://databases.about.com/od/specificproducts/l/aa3nf.htm
17. http://databases.about.com/cs/specificproducts/g/candidate.htm
18. http://databases.about.com/cs/specificproducts/g/candidate.htm
19. http://databases.about.com/cs/specificproducts/g/bcnf.htm
20. http://databases.about.com/od/specificproducts/g/multivalued-dependency.htm
21. http://databases.about.com/od/specificproducts/a/Should-I-Normalize-My-Database.htm
22. http://databases.about.com/od/specificproducts/a/normalization.htm
23. http://databases.about.com/od/specificproducts/l/aa1nf.htm
24. http://databases.about.com/od/specificproducts/a/2nf.htm
25. http://databases.about.com/od/specificproducts/l/aa3nf.htm
26. http://databases.about.com/gi/pages/mmail.htm