

# DB2 Normalization Reference Summary

SYS-ED Commerce & Programming courses include:

- Basics
- Installation & Support
- Web Authoring/Implementation
- Groupware
- WebSphere
- MQSeries
- DB2 UDB
- Programming Languages
- Internet Programming
- E-Commerce and Website Development

To see a list of courses available, go to our web site at [www.sysed.com](http://www.sysed.com).

*Get on the Fast Track!*



TM

## Normalization Reference Summary

Normalization helps eliminate redundancies and inconsistencies in table data. It is the process of reducing tables to a set of columns where all the non-key columns depend on the primary key column.

If this is not the case, the data can become inconsistent during updates.

## Form Description Overview

First	At each row and column position in the table, there exists one value, never a set of values.
Second	Each column that is not part of the key is dependent upon the key.
Third	Each non-key column is independent of other non-key columns, and is dependent only upon the key.
Fourth	No row contains two or more independent multi-valued facts about an entity.

## First Normal Form

A table is in first normal form if there is only one value, never a set of values, in each cell. A table that is in first normal form does not necessarily satisfy the criteria for higher normal forms.

## Second Normal Form

A table is in second normal form if each column that is not part of the key is dependent upon the entire key. Second normal form is violated when a non-key column is dependent upon part of a composite key.

The problems with violating the second normal form are:

- Data is repeated in every record.
- If the data needs to change, every row referring to the foreign key must be updated.
- Because of this redundancy, the data might become inconsistent, with different records.

There is a performance consideration in having the two tables in second normal form. Applications that produce reports using data from two tables must join both tables to retrieve the relevant information.



### Third Normal Form

A table is in third normal form if each non-key column is independent of other non-key columns, and is dependent only on the key.

The table can be normalized by creating a new table, with columns for the two tables.

Certain types of updates are easier; since only the new table needs to be updated.

An SQL query is more complex to write, because it requires joining the two tables.

It will probably also take longer to run than a query on a single table.

### Fourth Normal Form

A table is in fourth normal form if no row contains two or more independent multi-valued facts about an entity. A table is not in fourth normal form if it represents both relationships.

If, however, the attributes are interdependent, the table should not be split.

A good strategy when designing a database is to arrange all data in tables that are in fourth normal form, and then to decide whether the results give you an acceptable level of performance. If they do not, you can rearrange the data in tables that are in third normal form, and then reassess performance.