Second normal form

From Wikipedia, the free encyclopedia

Second normal form (**2NF**) is a normal form used in database normalization. 2NF was originally defined by E.F. Codd[1] in 1971. A table that is in first normal form (1NF) must meet additional criteria if it is to qualify for second normal form. Specifically: a 1NF table is in 2NF if and only if, given any candidate key and any attribute that is not a constituent of a candidate key, the non-key attribute depends upon the whole of the candidate key rather than just a part of it.

In slightly more formal terms: a 1NF table is in 2NF if and only if none of its non-prime attributes are functionally dependent on a part (proper subset) of a candidate key. (A non-prime attribute is one that does not belong to any candidate key.)

Note that when a 1NF table has no composite candidate keys (candidate keys consisting of more than one attribute), the table is automatically in 2NF.

Contents

- 1 Example
- 2 2NF and candidate keys
- 3 References
- 4 Further reading
- 5 External links

Example

Consider a table describing employees' skills:

Employees' Skills

Employee	Skill	Current Work Location
Jones	Typing	114 Main Street
Jones	Shorthand	114 Main Street
Jones	Whittling	114 Main Street
Roberts	Light Cleaning	73 Industrial Way
Ellis	Alchemy	73 Industrial Way
Ellis	Juggling	73 Industrial Way
Harrison	Light Cleaning	73 Industrial Way

The table's only candidate key is {Employee, Skill}.

The remaining attribute, Current Work Location, is dependent on only part of the candidate key, namely Employee. Therefore the table is not in 2NF. Note the redundancy in the way Current Work Locations are represented: we are told three times that Jones works at 114 Main Street, and twice that Ellis works at 73 Industrial Way. This redundancy makes the table vulnerable to update anomalies: it is, for example, possible to update Jones' work location on his "Typing" and "Shorthand" records and not update his "Whittling" record. The resulting data would imply contradictory answers to the question "What is Jones' current work location?"

A 2NF alternative to this design would represent the same information in two tables:

Employees

Employee	Current Work Location
Jones	114 Main Street
Roberts	73 Industrial Way
Ellis	73 Industrial Way
Harrison	73 Industrial Way

Employees' Skills

Employee	Skill
Jones	Typing
Jones	Shorthand
Jones	Whittling
Roberts	Light Cleaning
Ellis	Alchemy
Ellis	Juggling
Harrison	Light Cleaning

Update anomalies cannot occur in these tables, which are both in 2NF.

Not all 2NF tables are free from update anomalies, however. An example of a 2NF table which suffers from update anomalies is:

Tournament Winners

Tournament	Year	Winner	Winner Date of Birth
Des Moines Masters	1998	Chip Masterson	14 March 1977
Indiana Invitational	1998	Al Fredrickson	21 July 1975
Cleveland Open	1999	Bob Albertson	28 September 1968
Des Moines Masters	1999	Al Fredrickson	21 July 1975
Indiana Invitational	1999	Chip Masterson	14 March 1977

Even though Winner and Winner Date of Birth are determined by the whole key {Tournament, Year} and not part of it, particular Winner / Winner Date of Birth combinations are shown redundantly on multiple records. This problem is addressed by third normal form (3NF).

2NF and candidate keys

A table for which there are no partial functional dependencies on the primary key is typically, but not always, in 2NF. In addition to the primary key, the table may contain other candidate keys; it is necessary to establish that no non-prime attributes have part-key dependencies on **any** of these candidate keys.

Multiple candidate keys occur in the following table:

Electric Toothbrush Models

Manufacturer	Model	Model Full Name	Manufacturer Country
Forte	X-Prime	Forte X-Prime	Italy
Forte	Ultraclean	Forte Ultraclean	Italy
Dent-o-Fresh	EZBrush	Dent-o-Fresh EZBrush	USA
Kobayashi	ST-60	Kobayashi ST-60	Japan
Hoch	Toothmaster	Hoch Toothmaster	Germany
Hoch	Contender	Hoch Contender	Germany

Even if the designer has specified the primary key as {Model Full Name}, the table is not in 2NF. {Manufacturer, Model} is also a candidate key, and Manufacturer Country is dependent on a proper subset of it: Manufacturer.

References

1. ^ Codd, E.F. "Further Normalization of the Data Base Relational Model." (Presented at Courant Computer Science Symposia Series 6, "Data Base Systems," New York City, May 24th-25th, 1971.) IBM Research Report RJ909 (August 31st, 1971). Republished in Randall J. Rustin (ed.), *Data Base Systems: Courant Computer Science Symposia Series 6*. Prentice-Hall, 1972.

Further reading

- Litt's Tips: Normalization (http://www.troubleshooters.com/littstip/ltnorm.html)
- Rules Of Data Normalization (http://www.datamodel.org/NormalizationRules.html)
- Date, C. J., & Lorentzos, N., & Darwen, H. (2002). *Temporal Data & the Relational Model (http://www.elsevier.com/wps/product/cws_home/680662)* (1st ed.). Morgan Kaufmann. ISBN 1-55860-855-9.
- Date, C. J. (1999), *An Introduction to Database Systems (http://www.aw-bc.com/catalog/academic/product/0,1144,0321197844,00.html)* (8th ed.). Addison-Wesley Longman. ISBN 0-321-19784-4.
- Kent, W. (1983) *A Simple Guide to Five Normal Forms in Relational Database Theory (http://www.bkent.net/Doc/simple5.htm)*, Communications of the ACM, vol. 26, pp. 120-125
- Date, C.J., & Darwen, H., & Pascal, F. *Database Debunkings (http://www.dbdebunk.com/)*

External links

- Database Normalization Basics (http://databases.about.com/od/specificproducts/a/normalization.htm) by Mike Chapple (About.com)
- An Introduction to Database Normalization (http://dev.mysql.com/tech-resources/articles/intro-to-normalization.html) by Mike Hillyer.
- Normalization (http://www.utexas.edu/its/windows/database/datamodeling/rm/rm7.html) by ITS, University of Texas.
- A tutorial on the first 3 normal forms (http://phlonx.com/resources/nf3/) by Fred Coulson
- Free PDF poster available (http://www.marcrettig.com/poster/) by Marc Rettig
- Description of the database normalization basics (http://support.microsoft.com/kb/283878) by Microsoft

Topics in Database normalization

First normal form | **Second normal form** | Third normal form
Boyce-Codd normal form | Fourth normal form | Fifth normal form | Domain/key normal form |
Sixth normal form
Denormalization

Retrieved from "http://en.wikipedia.org/wiki/Second_normal_form"

Category: Database normalization

- This page was last modified 14:17, 29 November 2007.
- All text is available under the terms of the GNU Free Documentation License. (See **Copyrights** for details.) Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a U.S. registered 501(c)(3) tax-deductible nonprofit charity.