

Beispiel einer Normalisierung nach E. F. Codd

Normalisierung nach E. F. Codd

Ausgangssituation

Ein Verein will den Verleih von nur kurzzeitig benötigtem Spielzeug organisieren (Ludothek). Die Verleihvorgänge sollen in einer Datenbank gespeichert werden.

Schritt 0 (Vorbereitung)

Es sind alle Daten aufzulisten, die gespeichert werden sollen. Die Auflistung ist mit Beispieldaten zu versehen. Folgende Daten liegen nach der Datensammlung vor:

Spielmittelverleih (unnormalisiert)

Kunden- nr	Kundenname	Kundenanschrift	Verleih- nr	Verleih- datum	SpielNr	Spielbezeichnung	Rückgabe am
101	Müller, Fritz	12345 Berlin, Mustergasse 4	5	01.08.2008	21, 15	Sitzball, Bausteine	03.08.2008
101	Müller, Fritz	12345 Berlin, Mustergasse 4	5	01.08.2008	37	Traktor	
102	Meier, Frieda	23456 Hamburg, Webergasse 1	1	05.07.2008	15	Bausteine	15.07.2008
102	Meier, Frieda	23456 Hamburg, Webergasse 1	6	04.08.2008	21	Sitzball	
103	Lehmann, Max	12345 Berlin, Fallweg 1	3	10.07.2008	21, 37	Sitzball, Traktor	20.07.2008

Normalisierung nach E. F. Codd

Schritt 1 – 1. Normalform (1NF)

Die erste Normalform legt fest, dass Tabellen nur aus atomaren Werten bestehen und es keine Wiederholgruppen geben darf.

Spielmittelverleih (unnormalisiert)

Kunden- nr	Kundenname	Kundenanschrift	Verleih- nr	Verleih- datum	SpielNr	Spielbezeichnung	Rückgabe am
101	Müller, Fritz	12345 Berlin, Mustergasse 4	5	01.08.2008	21, 15	Sitzball, Bausteine	03.08.2008
101	Müller, Fritz	12345 Berlin, Mustergasse 4	5	01.08.2008	37	Traktor	
102	Meier, Frieda	23456 Hamburg, Webergasse 1	1	05.07.2008	15	Bausteine	15.07.2008
102	Meier, Frieda	23456 Hamburg, Webergasse 1	6	04.08.2008	21	Sitzball	
103	Lehmann, Max	12345 Berlin, Fallweg 1	3	10.07.2008	21, 37	Sitzball, Traktor	20.07.2008

Nicht atomar

Kundenname und Anschrift besteht jeweils aus mehreren unterschiedlichen Werten

Wiederholgruppen

Im Kreuzungspunkt von Zeile und Spalte sind mehrere Werte gespeichert.

Normalisierung nach E. F. Codd

Schritt 1 – 1. Normalform (1NF)

Die Attribute Kundenname und Kundenanschrift werden atomarisiert.

Spielmittelverleih (atomarisiert)

Kunden-nr	Nach-name	Vor-name	PLZ	Ort	Straße	Haus-nr	Verleih-nr	Verleih-datum	SpielNr	Spielbezeichnung	Rückgabe-am
101	Müller	Fritz	12345	Berlin	Mustergasse	4	5	01.08.2008	21, 15	Sitzball, Bausteine	03.08.2008
101	Müller	Fritz	12345	Berlin	Mustergasse	4	5	01.08.2008	37	Traktor	
102	Meier	Frieda	23456	Hamburg	Webergasse	1	1	05.07.2008	15	Bausteine	15.07.2008
102	Meier	Frieda	23456	Hamburg	Webergasse	1	6	04.08.2008	21	Sitzball	
103	Lehmann	Max	12345	Berlin	Fallweg	1	3	10.07.2008	21, 37	Sitzball, Traktor	20.07.2008

Wiederholgruppen

Im Kreuzungspunkt von Zeile und Spalte sind mehrere Werte gespeichert. Diese sind nun zu beseitigen.

Normalisierung nach E. F. Codd

Schritt 1 – 1. Normalform (1NF)

Spielmittelverleih (atomarisiert)

Die Wiederholgruppen Spielnr und Spielbezeichnung sind aufzulösen.

Kunden-nr	Nach-name	Vor-name	PLZ	Ort	Straße	Haus-nr	Verleih-nr	Verleih-datum	Spielnr	Spielbezeichnung	Rückgabe-am
101	Müller	Fritz	12345	Berlin	Mustergasse	4	5	01.08.2008	21, 15	Sitzball, Bausteine	03.08.2008
101	Müller	Fritz	12345	Berlin	Mustergasse	4	5	01.08.2008	37	Traktor	
102	Meier	Frieda	23456	Hamburg	Webergasse	1	1	05.07.2008	15	Bausteine	15.07.2008
102	Meier	Frieda	23456	Hamburg	Webergasse	1	6	04.08.2008	21	Sitzball	
103	Lehmann	Max	12345	Berlin	Fallweg	1	3	10.07.2008	21, 37	Sitzball, Traktor	20.07.2008

Spielmittelverleih (1NF)

Kunden-nr	Nach-name	Vor-name	PLZ	Ort	Straße	Haus-nr	Verleih-nr	Verleih-datum	Spiel-nr	Spielbezeichnung	Rückgabe-am
101	Müller	Fritz	12345	Berlin	Mustergasse	4	5	01.08.2008	21	Sitzball	03.08.2008
101	Müller	Fritz	12345	Berlin	Mustergasse	4	5	01.08.2008	15	Bausteine	03.08.2008
101	Müller	Fritz	12345	Berlin	Mustergasse	4	5	01.08.2008	37	Traktor	
102	Meier	Frieda	23456	Hamburg	Webergasse	1	1	05.07.2008	15	Bausteine	15.07.2008
102	Meier	Frieda	23456	Hamburg	Webergasse	1	6	04.08.2008	21	Sitzball	
103	Lehmann	Max	12345	Berlin	Fallweg	1	3	10.07.2008	21	Sitzball	20.07.2008
103	Lehmann	Max	12345	Berlin	Fallweg	1	3	10.07.2008	37	Traktor	20.07.2008

Nunmehr sind alle Felder atomar. Es gibt keine Wiederholgruppen → 1NF liegt vor

Normalisierung nach E. F. Codd

Schritt 2 – 2. Normalform (2NF)

Die zu speichernden Daten sind so in Tabellen aufzuteilen, dass jedes Nichtschlüsselattribut voll von einem Primärschlüssel abhängig ist.

Spielmitteilverleih (1NF)

Kunden- nr	Nach- name	Vor- name	PLZ	Ort	Straße	Haus- nr	Verleih- nr	Verleih- datum	Spiel- nr	Spielbe- zeichnung	Rückgabe am
101	Müller	Fritz	12345	Berlin	Mustergasse	4	5	01.08.2008	21	Sitzball	03.08.2008
101	Müller	Fritz	12345	Berlin	Mustergasse	4	5	01.08.2008	15	Bausteine	03.08.2008
101	Müller	Fritz	12345	Berlin	Mustergasse	4	5	01.08.2008	37	Traktor	
102	Meier	Frieda	23456	Hamburg	Webergasse	1	1	05.07.2008	15	Bausteine	15.07.2008
102	Meier	Frieda	23456	Hamburg	Webergasse	1	6	04.08.2008	21	Sitzball	
103	Lehmann	Max	12345	Berlin	Fallweg	1	3	10.07.2008	21	Sitzball	20.07.2008
103	Lehmann	Max	12345	Berlin	Fallweg	1	3	10.07.2008	37	Traktor	20.07.2008

Welche Attribute sind **Primärschlüssel**?

Welche Felder sind **Nichtschlüsselattribute**?


Normalisierung nach E. F. Codd

Schritt 2 – 2. Normalform (2NF)

Es ist zu überprüfen, ob es redundante Informationen gibt, die alle in gleicher Weise von einem Primärschlüssel abhängig sind.

Spielmittelverleih (1NF)

Kunden-nr	Nach-name	Vor-name	PLZ	Ort	Straße	Haus-nr	Verleih-nr	Verleih-datum	Spiel-nr	Spielbe-zeichnung	Rückgabe-am
101	Müller	Fritz	12345	Berlin	Mustergasse	4	5	01.08.2008	21	Sitzball	03.08.2008
101	Müller	Fritz	12345	Berlin	Mustergasse	4	5	01.08.2008	15	Bausteine	03.08.2008
101	Müller	Fritz	12345	Berlin	Mustergasse	4	5	01.08.2008	37	Traktor	
102	Meier	Frieda	23456	Hamburg	Webergasse	1	1	05.07.2008	15	Bausteine	15.07.2008
102	Meier	Frieda	23456	Hamburg	Webergasse	1	6	04.08.2008	21	Sitzball	
103	Lehmann	Max	12345	Berlin	Fallweg	1	3	10.07.2008	21	Sitzball	20.07.2008
103	Lehmann	Max	12345	Berlin	Fallweg	1	3	10.07.2008	37	Traktor	20.07.2008

Die Angaben Nachname, Vorname, PLZ, Ort Straße, Hausnr sind immer zugehörig zu einer Kundennr. Deshalb werden diese Felder als Kundentabelle ausgelagert: 

Normalisierung nach E. F. Codd

Schritt 2 – 2. Normalform (2NF)

Die Kundeninformationen werden ausgelagert:
Rest aus 1NF

Kunden- nr		Verleih- nr	Verleih- datum	Spiel- nr	Spielbe- zeichnung	Rückgabe am
101		5	01.08.2008	21	Sitzball	03.08.2008
101		5	01.08.2008	15	Bausteine	03.08.2008
101		5	01.08.2008	37	Traktor	
102		1	05.07.2008	15	Bausteine	15.07.2008
102		6	04.08.2008	21	Sitzball	
103		3	10.07.2008	21	Sitzball	20.07.2008
103		3	10.07.2008	37	Traktor	20.07.2008

Tabelle Kunden

Kunden- nr	Nach- name	Vor- name	PLZ	Ort	Straße	Haus-nr
101	Müller	Fritz	12345	Berlin	Mustergasse	4
102	Meier	Frieda	23456	Hamburg	Webergasse	1
103	Lehmann	Max	12345	Berlin	Fallweg	1

Der Rest wird weiter
bearbeitet:

Normalisierung nach E. F. Codd

Schritt 2 – 2. Normalform (2NF)

Der Rest aus der 1NF besteht aus folgenden Informationen:

Rest aus 1NF

Kunden- nr	Verleih- nr	Verleih- datum	Spiel- nr	Spielbe- zeichnung	Rückgabe am
101	5	01.08.2008	21	Sitzball	03.08.2008
101	5	01.08.2008	15	Bausteine	03.08.2008
101	5	01.08.2008	37	Traktor	
102	1	05.07.2008	15	Bausteine	15.07.2008
102	6	04.08.2008	21	Sitzball	
103	3	10.07.2008	21	Sitzball	20.07.2008
103	3	10.07.2008	37	Traktor	20.07.2008

Wieder ist zu überprüfen, ob es redundante Informationen gibt, die alle von ein und demselben Primärschlüssel abhängig sind.

Die Spielbezeichnung ist abhängig vom Primärschlüssel Spielnr. Deshalb wird eine Tabelle Spiele ausgelagert:

Normalisierung nach E. F. Codd

Schritt 2 – 2. Normalform (2NF)

Die Informationen zu den Spielen werden ausgelagert:

Rest aus 1NF

Kunden- nr	Verleih- nr	Verleih- datum	Spiel- nr		Rückgabe am
101	5	01.08.2008	21		03.08.2008
101	5	01.08.2008	15		03.08.2008
101	5	01.08.2008	37		
102	1	05.07.2008	15		15.07.2008
102	6	04.08.2008	21		
103	3	10.07.2008	21		20.07.2008
103	3	10.07.2008	37		20.07.2008

Der Rest wird weiter
bearbeitet:

Tabelle Spiele

Spiel- nr	Spielbe- zeichnung
21	Sitzball
15	Bausteine
37	Traktor

Normalisierung nach E. F. Codd

Schritt 2 – 2. Normalform (2NF)

Der Rest aus der 1NF besteht aus folgenden Informationen:

Rest aus 1NF

Kunden- nr	Verleih- nr	Verleih- datum	Spiel- nr	Rückgabe am
101	5	01.08.2008	21	03.08.2008
101	5	01.08.2008	15	03.08.2008
101	5	01.08.2008	37	
102	1	05.07.2008	15	15.07.2008
102	6	04.08.2008	21	
103	3	10.07.2008	21	20.07.2008
103	3	10.07.2008	37	20.07.2008

Wieder ist zu überprüfen, ob es redundante Informationen gibt, die alle von ein und demselben Primärschlüssel abhängig sind.

Zur Verleihrnr gehören eindeutig die Kundennr und das Verleihdatum. Deshalb wird nun eine Tabelle Verleih ausgelagert:

Normalisierung nach E. F. Codd

Schritt 2 – 2. Normalform (2NF)

Die Informationen zum Verleih werden ausgelagert:
Rest aus 1NF

Verleih- nr		Spiel- nr	Rückgabe am
5		21	03.08.2008
5		15	03.08.2008
5		37	
1		15	15.07.2008
6		21	
3		21	20.07.2008
3		37	20.07.2008

Der Rest wird weiter bearbeitet:

Tabelle Verleih

Kunden- nr	Verleih- nr	Verleih- datum
101	5	01.08.2008
102	1	05.07.2008
102	6	04.08.2008
103	3	10.07.2008

Schritt 2 – 2. Normalform (2NF)

Der Rest aus der 1NF besteht aus folgenden Informationen:

Rest aus 1NF

Verleih-nr	Spiel-nr	Rückgabe am
5	21	03.08.2008
5	15	03.08.2008
5	37	
1	15	15.07.2008
6	21	
3	21	20.07.2008
3	37	20.07.2008

Der Rest enthält keine redundanten Informationen. Deshalb kann nichts mehr ausgelagert werden. Der Rest aus 1NF wird als Tabelle Verleihdetails bezeichnet.

Tabelle Verleihdetails

Verleih-nr	Spiel-nr	Rückgabe am
5	21	03.08.2008
5	15	03.08.2008
5	37	
1	15	15.07.2008
6	21	
3	21	20.07.2008
3	37	20.07.2008

Normalisierung nach E. F. Codd

Schritt 2 – 2. Normalform (2NF)

Das Endergebnis der 2NF besteht damit aus den 4 Tabellen

Kunden						
Kunden-nr	Nach-name	Vor-name	PLZ	Ort	Straße	Haus-nr
101	Müller	Fritz	12345	Berlin	Mustergasse	4
102	Meier	Frieda	23456	Hamburg	Webergasse	1
103	Lehmann	Max	12345	Berlin	Fallweg	1

Spiele	
Spiel-nr	Spielbe-zeichnung
15	Bausteine
21	Sitzball
37	Traktor

Verleih		
Verleih-nr	KNr	Verleih-datum
5	101	01.08.2008
1	102	05.07.2008
6	102	04.08.2008
3	103	10.07.2008

Verleihdetails		
Verleih-nr	Spiel-nr	Rückgabe-am
5	21	03.08.2008
5	15	03.08.2008
5	37	
1	15	15.07.2008
6	21	
3	21	20.07.2008
3	37	20.07.2008

Normalisierung nach E. F. Codd

Schritt 3 – 3. Normalform (3NF)

In der 3NF darf es keine transitiven Abhängigkeiten geben.

Kunden						
Kunden-nr	Nach-name	Vor-name	PLZ	Ort	Straße	Haus-nr
101	Müller	Fritz	12345	Berlin	Mustergasse	4
102	Meier	Frieda	23456	Hamburg	Webergasse	1
103	Lehmann	Max	12345	Berlin	Fallweg	1

Verleih		
Verleih-nr	KNr	Verleih-datum
5	101	01.08.2008
1	102	05.07.2008
6	102	04.08.2008
3	103	10.07.2008

Verleihdetails		
Verleih-nr	Spiel-nr	Rückgabe am
5	21	03.08.2008
5	15	03.08.2008
5	37	
1	15	15.07.2008
6	21	
3	21	20.07.2008
3	37	20.07.2008

Spiele	
Spiel-nr	Spielbe-zeichnung
15	Bausteine
21	Sitzball
37	Traktor

Nachdem alle Nichtschlüsselattribute eindeutig abhängig sind von einem Primärschlüssel ist innerhalb der Tabellen zu überprüfen, ob transitive Abhängigkeiten vorhanden sind. Diese transitiven Abhängigkeiten sind in der 3NF zu beseitigen.

Normalisierung nach E. F. Codd

Schritt 3 – 3. Normalform (3NF)

Kunden						
Kunden-nr	Nach-name	Vor-name	PLZ	Ort	Straße	Haus-nr
101	Müller	Fritz	2345	Berlin	Mustergasse	4
102	Meier	Frieda	23456	Hamburg	Webergasse	1
103	Lehmann	Max	2345	Berlin	Fallweg	1

Verleih		
Verleih-nr	KNr	Verleih-datum
5	101	01.08.2008
1	102	05.07.2008
6	102	04.08.2008
3	103	10.07.2008

Verleihdetails		
Verleih-nr	Spiel-nr	Rückgabe am
5	21	03.08.2008
5	15	03.08.2008
5	37	
1	15	15.07.2008
6	21	
3	21	20.07.2008
3	37	20.07.2008

Spiele	
Spiel-nr	Spielbezeichnung
15	Bausteine
21	Sitzball
37	Traktor

Überführung in 3NF

Spielmittelverleih (3NF)

Kunden					
Kunden-nr	Nach-name	Vor-name	PLZ	Straße	Haus-nr
101	Müller	Fritz	12345	Mustergasse	4
102	Meier	Frieda	23456	Webergasse	1
103	Lehmann	Max	12345	Fallweg	1

Verleih		
Verleih-nr	KNr	Verleih-datum
5	101	01.08.2008
1	102	05.07.2008
6	102	04.08.2008
3	103	10.07.2008

Verleihdetails		
Verleih-nr	Spiel-nr	Rückgabe am
5	21	03.08.2008
5	15	03.08.2008
5	37	
1	15	15.07.2008
6	21	
3	21	20.07.2008
3	37	20.07.2008

Spiele	
Spiel-nr	Spielbezeichnung
15	Bausteine
21	Sitzball
37	Traktor

Orte	
PLZ	Ort
12345	Berlin
23456	Hamburg

Normalisierung nach E. F. Codd

Die 3NF ist die Form, in der Daten für die Speicherung in relationalen Datenbanken in der Regel aufbereitet werden. Auf dieser Basis lässt sich ein Datenmodell erstellen, das alle Anforderungen erfüllt.

Wird dieses Datenmodell in Access umgesetzt, sieht das Beziehungsfenster wie nachfolgend aus:

