

Normalformen

- Normalformen
 - 1 bis 5 und Boyce-Codd-Normalform/BCNF
- Unter anderem werden Normalformen dazu verwendet, um unerwünschte Abhängigkeiten bei DELETE-, UPDATE- und INSERT-Operationen (Anomalien) zu vermeiden.
- darüber hinaus spricht man noch von
 - Unnormalisierter Form: betrifft alle Datendefinitionen, die nicht mindestens der 1. Normalform genügen ;-))

Anomalien

In folgendem Beispiel entstehen folgende Anomalien:

<u>CD ID</u>	<u>CD-Titel</u>	<u>Jahr</u>	<u>Interpret</u>	<u>Track Nr</u>	<u>Song</u>
1	Let it be	1987	The Beatles	1	Two of us
1	Let it be	1987	The Beatles	2	Dig a Pony
	...				

Insert-Anomalie: Ein Interpret kann erst gespeichert werden, wenn er eine CD herausgebracht hat.

Delete-Anomalie: Eine CD kann nur gelöscht werden, wenn auch der Interpret gelöscht wird.

Update-Anomalie: Wenn sich der Name eines Interpreten ändert, müssen alle Einträge in der Tabelle geändert werden.

Erste Normalform (1. NF)

- Ein Relationstyp ist in der 1. Normalform, wenn **alle Attribute maximal einen Wert** haben. Am Kreuzungspunkt einer Spalte mit einer Reihe darf also maximal ein Datenwert stehen. **Attribute sind atomar.**
Beispiel: Mobilnummer, Festnetznummer usw. dürfen nicht in einer Spalte sein.
- Das Nichtvorhandensein von Daten ist zulässig.

Beispiel

Personendaten und zwar Nachname und Vorname sind zu speichern.

Lösungsidee

1. Entwurf

Tabelle PERSON

PID (Personal Identifier)	Name
1	Regina Martl
2	Jürgen Strutzenberger
3	Sevinc Dursun

Das Attribut NAME hat mehr als einen Wert je Datensatz und verletzt somit die 1. NF.

2. Entwurf – Verbesserung 1. Entwurf

Tabelle PERSON

PID	Vorname	Nachname
1	Regina	Martl
2	Jürgen	Strutzenberger
3	Sevinc	Dursun

Zweite Normalform

- Ein Relationstyp ist in der 2. Normalform, wenn er in der 1. Normalform ist und **jedes Attribut (ohne FK) von ALLEN Schlüsselkandidaten (Entities) funktional** abhängt.
- Schlüsselkandidaten sind jene Attribute, die einen Datensatz eindeutig ausweisen (dies kann EIN Attribut sein, z. B. Personalnummer, oder MEHRERE GEMEINSAM z. B. Personenummer UND Geburtsdatum = Sozialversicherungsnummer).
- Meist führt die Verletzung der 2. NF zu sogenannten Anomalien (siehe Beispiel).
- Relationstypen, die in der 1. Normalform sind, sind automatisch in der 2. Normalform, wenn ihr (erster) Primärschlüssel nicht zusammengesetzt ist (Ausnahme Zwischentabellen).

Beispiel

Ein Möbelhaus speichert alle Produkte inkl. Hersteller, Produktnummer und Typ.

Lösungsidee

1. Entwurf

Tabelle **PRODUKTHERSTELLER**

<u>HerstellerID</u>	Hersteller	<u>ProduktID</u>	Produkttype
1717	Moses	245613	Kasten
1718	Warter	987439	Sofa
1717	Moses	987439	Sofa

Als Schlüsselkandidaten gelten hier HerstellerID und ProduktID.

Überprüfung:

Schlüsselkandidat HerstellerID

mit Hersteller

1717 Moses
1718 Warter

mit Produkttype

1717 Kasten UND Sofa
1718 Sofa

Schlüsselkandidat ProduktID

mit Hersteller

245613 Moses
987439 Warter UND Moses

mit Produkttyp

245613 Kasten
987439 Sofa

Für ProduktID 987439 gibt es zwei Hersteller und für HerstellerID 1717 gibt es zwei Produkttypen. Daher 2. NF verletzt.

2. Entwurf

Tabelle **HERSTELLER**

<u>HerstellerID</u>	Hersteller
1717	Moses
1718	Warter

Tabelle **PRODUKTHERSTELLER**

<u>HerstellerID</u>	<u>ProduktID</u>
1717	245613
1718	987439
1717	987439

Tabelle **PRODUKT**

<u>ProduktID</u>	Produkttype
245613	Kasten
987439	Sofa

Dritte Normalform (3. NF)

- Die 3. Normalform ist erfüllt, wenn die 2. Normalform erfüllt ist und die **Nicht-Schlüssel-Attribute funktional unabhängig voneinander** sind. Null-Werte bei Unique Keys (zusammengesetzter PK, weilers FK) sind nicht erlaubt (Ausnahmen möglich).
- Transitive Abhängigkeit: wenn ein Attribut nicht nur durch alle Schlüssel, bzw. einen identifiziert werden kann, sondern auch durch einen einzelnen.
- Eine funktionale Abhängigkeit kann auch von einer Gruppe von Attributen bestehen.

Beispiel

Weiterführendes Beispiel zu Beispiel aus 2.NF. Zu den Möbeln wird auch das Grundmaterial mit Materialcode gespeichert.

1. Entwurf

Tabelle **HERSTELLER**

<u>HerstellerID</u>	Hersteller
1717	Moses
1718	Warter

Tabelle **PRODUKTHERSTELLER**

<u>HerstellerID</u>	ProduktID
1717	245613
1718	987439
1717	987439
1717	27

Tabelle **PRODUKT**

<u>ProduktID</u>	Produkttype	Materialcode	Material
245613	Kasten	XXZF	Holz
987439	Sofa	4FER	Stoff
272712	Tisch	XXZF	Holz

Nicht-Schlüsselattribute in der Tabelle Produkt sind: Produkttype, Materialcode, Material

Überprüfung:

Attribut Produkttype

mit Materialcode

Kasten	XXZF
Sofa	4FER
Tisch	XXZF

mit Material

Kasten	Holz
Sofa	Stoff
Tisch	Holz

Attribut Materialcode
mit Produkttype

XXZF	Kasten
4FER	Sofa
XXZF	Tisch

mit Material

XXZF	Holz
4FER	Sofa
XXZF	Holz

Für Holz kommt immer der Materialcode XXZF und auch umgekehrt. Das Material ist nicht von der ProduktID abhängig. Daher Verletzung der 3. NF.

1. Entwurf

Tabelle **HERSTELLER**

2. HerstellerID	Hersteller
1717	Moses
1718	Warter

Tabelle **PRODUKTHERSTELLER**

HerstellerID	ProduktID
1717	245613
1718	987439
1717	987439
1717	272712

Tabelle **PRODUKT**

ProduktID	Produkttype
245613	Kasten
987439	Sofa
272712	Tisch

Tabelle **MATERIAL**

Materialcode	Material
XXZF	Holz
4FER	Stoff

Tabelle **PRODUKTMATERIAL**

Materialcode	ProduktID
XXZF	245613
4FER	987439
XXZF	271712