

# Übung- Normalisierung- Tabellen in die 3.NF überführen

## Übung 1

CD\_Lied

CD_ID	Album	Gründungsjahr	Erscheinungsjahr	Titelliste
4711	Anastacia – Not That Kind	1999	2000	{1. Not That Kind, 2. I'm Outta Love, 3. Cowboys & Kisses}
4712	Pink Floyd – Wish You Were Here	1965	1975	{1. Shine On You Crazy Diamond}
4713	Anastacia – Freak of Nature	1999	2001	{1. Paid my Dues}

## Übung 2

P#	P_Name	Abt#	Abt-Name	Pj#	Pj-Name	Pj-Std
101	Müller	1	Motoren	11, 12	A, B	60, 40
102	Meier	2	Karosserie	13	C	100
103	Krause	2	Karosserie	11, 12, 13	A, B, C	20, 50, 30
104	Schmidt	1	Motoren	11, 13	A, C	80, 20

P#:Personal Nummer  
Pj#:Projekt Nummer

CD_Lied				
CD_ID	Album	Gründungsjahr	Erscheinungsjahr	Titelliste
4711	Anastacia – Not That Kind	1999	2000	{1. Not That Kind, 2. I'm Outta Love, 3. Cowboys & Kisses}
4712	Pink Floyd – Wish You Were Here	1965	1975	{1. Shine On You Crazy Diamond}
4713	Anastacia – Freak of Nature	1999	2001	{1. Paid my Dues}

- Das Feld Album enthält die Attributwertebereiche Interpret und Albumtitel.
- Das Feld Titelliste enthält eine Menge von Titeln.

Dadurch hat man ohne Aufspaltung folgende Probleme bei Abfragen:

- Zur Sortierung nach Albumtitel muss das Feld Album in Interpret und Albumtitel aufgeteilt werden.
- Die Titel können (mit einfachen Mitteln) nur alle gleichzeitig als Titelliste oder gar nicht dargestellt werden.

## 1.NF

CD_Lied						
CD_ID	Albumtitel	Interpret	Gründungsjahr	Erscheinungsjahr	Track	Titel
4711	Not That Kind	Anastacia	1999	2000	1	Not That Kind
4711	Not That Kind	Anastacia	1999	2000	2	I'm Outta Love
4711	Not That Kind	Anastacia	1999	2000	3	Cowboys & Kisses
4712	Wish You Were Here	Pink Floyd	1965	1975	1	Shine On You Crazy Diamond
4713	Freak of Nature	Anastacia	1999	2001	1	Paid my Dues

Die Attributwertebereiche werden in atomare Attributwertebereiche aufgespalten:

- Das Feld Album wird in die Felder Albumtitel und Interpret gespalten.
- Das Feld Titelliste wird in die Felder Track und Titel gespalten sowie auf mehrere Datensätze aufgeteilt.

Da jetzt jeder Attributwertebereich atomar ist sowie die Tabelle einen eindeutigen Primärschlüssel (Verbundschlüssel aus den Spalten CD\_ID und Track) besitzt, befindet sich die Relation in 1NF.

## 2. NF

### 2NF verletzt

CD\_Lied

CD_ID	Albumtitel	Interpret	Gründungsjahr	Erscheinungsjahr	Track	Titel
4711	Not That Kind	Anastacia	1999	2000	1	Not That Kind
4711	Not That Kind	Anastacia	1999	2000	2	I'm Outta Love
4711	Not That Kind	Anastacia	1999	2000	3	Cowboys & Kisses
4712	Wish You Were Here	Pink Floyd	1965	1975	1	Shine On You Crazy Diamond
4713	Freak of Nature	Anastacia	1999	2001	1	Paid my Dues

- Der Primärschlüssel der Relation ist aus den Feldern CD\_ID und Track zusammengesetzt. (Grundsätzlich darf ein Primärschlüssel aus mehreren Attributen bestehen, jedoch entsteht daraus im genannten Beispiel ein Konflikt.)
- Die Felder Albumtitel, Interpret und Erscheinungsjahr sind vom Feld CD\_ID abhängig, aber nicht vom Feld Track. Dieser (Punkt 2) verletzt die 2. Normalform, da die drei nicht-primären Attribute nicht nur von einem Teil des Schlüssels (hier CD\_ID) abhängen dürfen. Wäre der Schlüssel nicht zusammengesetzt (siehe Punkt 1), so könnte dies nicht passieren.

## 2.NF

### Probleme, die sich daraus ergeben

Die Informationen aus diesen drei Feldern sind, wie am Beispiel der CD Not That Kind zu erkennen, mehrfach vorhanden, d. h. redundant. Dadurch besteht die Gefahr, dass die Integrität der Daten verletzt wird. So könnte man den Albumtitel für das Lied Not That Kind in I Don't Mind ändern, ohne jedoch die entsprechenden Einträge für die Titel I'm Outta Love und Cowboys & Kisses zu ändern (Update-Anomalie).

CD\_Lied (inkonsistent)

CD_ID	Albumtitel	Interpret	Gründungsjahr	Erscheinungsjahr	Track	Titel
4711	I Don't Mind	Anastacia	1999	2000	1	Not That Kind
4711	Not That Kind	Anastacia	1999	2000	2	I'm Outta Love
4711	Not That Kind	Anastacia	1999	2000	3	Cowboys & Kisses
4712	Wish You Were Here	Pink Floyd	1965	1975	1	Shine On You Crazy Diamond
4713	Freak of Nature	Anastacia	1999	2001	1	Paid my Dues

## 2.NF Lösung

CD					Lied		
CD_ID	Albumtitel	Interpret	Gründungsjahr	Erscheinungsjahr	CD_ID	Track	Titel
4711	Not That Kind	Anastacia	1999	2000	4711	1	Not That Kind
4712	Wish You Were Here	Pink Floyd	1965	1975	4711	2	I'm Outta Love
4713	Freak of Nature	Anastacia	1999	2001	4711	3	Cowboys & Kisses
					4712	1	Shine On You Crazy Diamond
					4713	1	Paid my Dues

Das Attribut CD\_ID aus der Tabelle Lied bezeichnet man als Fremdschlüssel, der auf den Primärschlüssel der Tabelle CD verweist. Zugleich stellen die Attribute CD\_ID und Track den zusammengesetzten Primärschlüssel der Tabelle Lied dar.

## Veletzt

CD				
CD_ID	Albumtitel	Interpret	Gründungsjahr	Erscheinungsjahr
4711	Not That Kind	Anastacia	1999	2000
4712	Wish You Were Here	Pink Floyd	1965	1975
4713	Freak of Nature	Anastacia	1999	2001

Das Problem ist hierbei wieder Datenredundanz. Wird zum Beispiel eine neue CD mit einem existierenden Interpreten eingeführt, so wird das Gründungsjahr redundant gespeichert.

## 3.NF Lösung

CD				Künstler		Lied		
CD_ID	Albumtitel	Interpret	Erscheinungsjahr	Interpret	Gründungsjahr	CD_ID	Track	Titel
4711	Not That Kind	Anastacia	2000	Anastacia	1999	4711	1	Not That Kind
4712	Wish You Were Here	Pink Floyd	1975	Pink Floyd	1965	4711	2	I'm Outta Love
4713	Freak of Nature	Anastacia	2001			4711	3	Cowboys & Kisses
						4712	1	Shine On You Crazy Diamond
						4713	1	Paid my Dues

Diese Lösung gilt nur, wenn man davon ausgeht, dass der Interpret weltweit eindeutig ist. Ansonsten müsste man eine synthetische ID in der Tabelle Künstler hinzufügen, die dann den Fremdschlüssel in der Tabelle CD stellt, wie folgt:



CD				Künstler			Lied		
CD_ID	Albumtitel	Interpret_ID	Erscheinungsjahr	Interpret_ID	Interpret	Gründungsjahr	CD_ID	Track	Titel
4711	Not That Kind	311	2000	311	Anastacia	1999	4711	1	Not That Kind
4712	Wish You Were Here	312	1975	312	Pink Floyd	1965	4711	2	I'm Outta Love
4713	Freak of Nature	311	2001				4711	3	Cowboys & Kisses
							4712	1	Shine On You Crazy Diamond
							4713	1	Paid my Dues

# 1. Übung: Normalisierung

- Eine Tabelle zur Verwaltung von Personalinformationen soll bis in die 3. Normalform überführt werden. Angelegt wurde eine „naive“ Tabelle „PERSONAL-PROJEKT“. Mitarbeiterinnen und Mitarbeiter können an mehreren Projekten beteiligt sein. Es wird jeweils notiert, wie viele Stunden sie in einem Projekt geleistet haben.

<u>P#</u>	P_Name	Abt#	Abt-Name	Pj#	Pj-Name	Pj-Std
101	Müller	1	Motoren	11, 12	A, B	60, 40
102	Meier	2	Karosserie	13	C	100
103	Krause	2	Karosserie	11, 12, 13	A, B, C	20, 50, 30
104	Schmidt	1	Motoren	11, 13	A, C	80, 20

# 1. Übung: Normalisierung

## Lösung: 1NF

- Eine Tabelle liegt in der ersten Normalform vor, wenn jeder Attributwert eine atomare, nicht weiter zerlegbare Dateneinheit ist.  
Alternativ: Eine Tabelle ist nicht in der 1NF, wenn Attribute mehrfach komplex in einer Spalte auftreten; d.h. 1NF ist eine Strukturierungsvorschrift.  
Rezept: Auslagern der nicht atomaren Attribute in verschiedene Zeilen oder Spalten oder eine eigene Tabelle.

<u>P#</u>	P_Name	Abt#	Abt-Name	Pj#	Pj-Name	Pj-Std
101	Müller	1	Motoren	11	A	60
101	Müller	1	Motoren	12	B	40
102	Meier	2	Karosserie	13	C	100
103	Krause	2	Karosserie	11	A	20
103	Krause	2	Karosserie	12	B	50
103	Krause	2	Karosserie	13	C	30
104	Schmidt	1	Motoren	11	A	80
104	Schmidt	1	Motoren	13	C	20

# 1. Übung: Normalisierung

## Lösung: 2NF

- Eine Tabelle liegt in der zweiten Normalform vor, wenn sie in der ersten Normalform ist und jedes Nichtschlüsselattribut voll funktional abhängig vom Primärschlüssel ist.  
Alternativ: Eine Tabelle ist nicht in der 2 NF, wenn Attribute von einem Teil des Schlüssels eindeutig identifiziert werden. Voraussetzung ist die 1NF.  
Rezept: Auslagern von Teilschlüsseln und zugehörigen Informationen in eigene Tabelle nach Sachgebieten; bzw. separaten Entitätstypen mit eigenem Schlüssel finden. Beim Auslagern durch entsprechende Beziehungen darauf achten, dass Informationen nicht verloren gehen.

- Tabelle: PERSONAL

P#	P-Name	Abt#	Abt-Name
101	Müller	1	Motoren
102	Meier	2	Karosserie
103	Krause	2	Karosserie
104	Schmidt	1	Motoren

- Tabelle: PROJEKT

Pj#	Pj-Name
11	A
12	B
13	C

- Tabelle: PERSONAL-PROJEKT

P#	Pj#	Pj-Std
101	11	60
101	12	40
102	13	100
103	11	20
103	12	50
103	13	30
104	11	80
104	13	20

# 1. Übung: Normalisierung

## Lösung: 3NF

- Eine Tabelle liegt in der 3NF vor, wenn sie sich in der 2NF befindet und jedes Nichtschlüsselattribut nicht transitiv abhängig vom Primärschlüssel ist.  
Alternativ: Eine Tabelle ist nicht in der 3NF, wenn Attribute von anderen NICHT-Schlüsselattributen identifiziert werden. Voraussetzung ist außerdem die 2NF.  
Rezept: Auslagern der „transitiv abhängigen“ Attribute in eigene Tabellen.

Tabelle: PERSONAL

P#	P-Name	Abt#
101	Müller	1
102	Meier	2
103	Krause	2
104	Schmidt	1

Tabelle: PROJEKT

Pj#	Pj-Name
11	A
12	B
13	C

Tabelle: PERSONAL-PROJEKT

P#	Pj#	Pj-Std
101	11	60
101	12	40
102	13	100
103	11	20
103	12	50
103	13	30
104	11	80
104	13	20

Tabelle: ABTEILUNG

Abt#	Abt-Name
1	Motoren
2	Karosserie