

## Normalisierung von Datenbanken

© D. Langhans

### Was bezweckt das? Dazu zwei Beispiele!

Artikel Nr	Artikelart	Artikel Bezeichnung	Verkaufspreis	Jahresumsatz	Kunden Nr
1000	Schreibtisch	Gabriele	1.365,00 €	250	101, 104, 105
2000	Scanner	Swift	2.198,00 €	150	104, 125
2001	Scanner	Akura	3.700,00 €	16	180, 202

Verkäufer	Wohnort	Kinder	Gehalt	Kunden	Kundenstandort
Müller	Köln	Lukas	2.000,00 €	Kronos AG	Ratingen
Müller	Köln	Karla	2.000,00 €	Kronos AG	Ratingen
Schmidt	Neuss		1.800,00 €	Bau KG	Köln
Kramer	Hilden		2.150,00 €	Heim AG	Hilden
Kramer	Hilden		2.150,00 €	Tierbau AB	Solingen

© D. Langhans

2

### Was heißt „Normalisierung“?

- Vernünftige Aufteilung der Daten in mehrere Tabellen
- Theorie: Normalformenlehre nach Codd (1970)
- Verschiedene Stufen: 5 in der Praxis werden 3 durchgeführt



**Edgar F. Codd**, amerikanischer Mathematiker und Wissenschaftler, schuf in den 1960er und 1970er Jahren bei seiner Forschungsarbeit am IBM Almaden Research Center in San Jose die Grundlagen der relationalen Datenbanken.

© D. Langhans

3

### 1. Normalform - Beispiel 1

- Erreicht, wenn alle Attribute nur atomare Werte enthalten

Artikel Nr	Artikelart	Artikel Bezeichnung	Verkaufspreis	Jahresumsatz	Kunden Nr
1000	Schreibtisch	Gabriele	1.365,00 €	250	101, 104, 105
2000	Scanner	Swift	2.198,00 €	150	104, 125
2001	Scanner	Akura	3.700,00 €	16	180, 202

- hier: nicht der Fall wegen „Kunden Nr“!

© D. Langhans

4

### 1. Normalform - Beispiel 1

Artikel Nr	Artikelart	Artikel Bezeichnung	Verkaufspreis	Jahresumsatz	Kunden Nr
1000	Schreibtisch	Gabriele	1.365,00 €	250	101
1000	Schreibtisch	Gabriele	1.365,00 €	250	104
1000	Schreibtisch	Gabriele	1.365,00 €	250	105
2000	Scanner	Swift	2.198,00 €	150	104
2000	Scanner	Swift	2.198,00 €	150	125
2001	Scanner	Akura	3.700,00 €	16	180
2001	Scanner	Akura	3.700,00 €	16	202

© D. Langhans

5

### 1. Normalform - Beispiel 2

- Erreicht, wenn alle Attribute nur atomare Werte enthalten

Artikelnummer	Artikeldaten	Jahresumsatz
2000	Scanner Swift, 2.198,-	150
2001	Scanner Acura, 3.700,-	16

- hier: nicht der Fall wegen „Artikeldaten“!

© D. Langhans

6

## 1. Normalform - Beispiel 2

- Erreicht, wenn alle Attribute nur atomare Werte enthalten

Artikelnr.	Artikelgruppe	Artikelnamen	Verkaufspreis	Jahresumsatz
2000	Scanner	Swift	2.198,-	150
2001	Scanner	Acura	3.700,-	16

© D. Langhank

7

## 2. Normalform

- Beseitigung unerwünschter Abhängigkeiten
- Vermeidung von Redundanzen
- Erreicht, wenn:
  - die 1. Normalform gegeben ist
  - alle Nichtschlüsselattribute voll von jedem Schlüssel (Primärschlüssel) abhängig sind
- Gibt es also einen **zusammengesetzten Schlüssel**, dürfen Attribute nicht nur von einem Teil dieses (zusammengesetzten) Schlüssels abhängig sein!

© D. Langhank

8

Mitarbeiter Nr.	Name	Wohnort	Kunden Nr.	Kunden	Kundenstandort	Umsatz
10	Müller	Köln	101	Kronos AG	Ratingen	10.000,00 €
10	Müller	Köln	102	Bau KG	Köln	15.000,00 €
11	Schmidt	Neuss	103	Grün AG	Köln	8.000,00 €
12	Kramer	Hilden	104	Hein AG	Hilden	9.000,00 €
12	Kramer	Hilden	105	Tiefbau AG	Solingen	12.000,00 €
12	Kramer	Hilden	103	Grün AG	Köln	5.000,00 €

© D. Langhank

9

Mitarbeiter Nr.	Name	Wohnort	Kunden Nr.	Kunden	Kundenstandort	Umsatz
10	Müller	Köln	101	Kronos AG	Ratingen	10.000,00 €
10	Müller	Köln	102	Bau KG	Köln	15.000,00 €
11	Schmidt	Neuss	103	Grün AG	Köln	8.000,00 €
12	Kramer	Hilden	104	Hein AG	Hilden	9.000,00 €
12	Kramer	Hilden	105	Tiefbau AG	Solingen	12.000,00 €
12	Kramer	Hilden	103	Grün AG	Köln	5.000,00 €

- „Mitarbeiter Nr.“ und „Kunden Nr.“ sind der zusammengesetzte Primärschlüssel
- „Wohnort“ und „Name“ ist von „Mitarbeiter Nr.“ abhängig
- „Kunden“ und „Kundenstandort“ ist von „Kunden“ abhängig
- Nur „Umsatz“ ist von beiden Schlüsseln abhängig
- Ergebnis: Redundanzen müssen eingegeben werden!
- z. B.:  
Mitarbeiter Kramer zieht von Hilden nach Langenfeld um  
Die Grün AG firmiert um in „Grün West AG“

© D. Langhank

10

## 2. Normalform: Aufteilen in 3 Tabellen

Mitarbeiter Nr.	Name	Wohnort
10	Müller	Köln
11	Schmidt	Neuss
12	Kramer	Hilden

  

Kunden Nr.	Kunden	Kundenstandort
101	Kronos AG	Ratingen
102	Bau KG	Köln
103	Grün AG	Köln
104	Hein AG	Hilden
105	Tiefbau AG	Solingen

  

Mitarbeiter Nr.	Kunden Nr.	Umsatz
10	101	10.000,00 €
10	102	15.000,00 €
11	103	8.000,00 €
12	104	9.000,00 €
12	105	12.000,00 €

© D. Langhank

11

## 3. Normalform

- Erreicht, wenn
  - die 2. Normalform gegeben ist
  - kein Nichtschlüsselattribut von irgendeinem Schlüssel transitiv abhängig ist.  
„Transitiv“ heißt „indirekt“.

© D. Langhank

12

Mitarbeiter Nr.	Name	Wohnort	Abteilungs-Nr.	Abteilung	Abteilungsleiter
10	Müller	Köln	5	Tiefbau	Kluge
11	Schmidt	Neuss	8	Hochbau	Weiss
12	Kramer	Hilden	6	Brückenbau	Böll

- „Abteilung“ und „Abteilungsleiter“ sind direkt von „Abteilungs-Nr.“ und transitiv (indirekt) von MitarbeiterNr. abhängig
- Erreichen der 3. Normalform: Zerlegen in zwei Tabellen

Mitarbeiter Nr.	Name	Wohnort	Abteilungs-Nr.
10	Müller	Köln	5
11	Schmidt	Neuss	8
12	Kramer	Hilden	6

Abteilungs-Nr.	Abteilung	Abteilungsleiter
5	Tiefbau	Kluge
8	Hochbau	Weiss
6	Brückenbau	Böll

© D. Langhans

13

## Denormalisierung

- Praktische Überlegungen sprechen manchmal gegen eine vollständige Normalisierung: zur Vermeidung sehr vieler Tabellen
- Beispiel:  
bei Tabellen, die Adressdaten aufnehmen, werden oft die Postleitzahl und der Ort NICHT in einer eigenen Tabelle aufgenommen - was gegen die 3. Normalform verstößt!
- Fazit:
  - die Normalisierung ist kein Dogma (Zwang)
  - aber ein analytisches Instrument, um Tabellenstrukturen auf ihre Qualität hin zu überprüfen

© D. Langhans

14