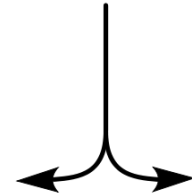


TBL_AdressenAlles
Firma
Strasse
PLZ
Ort

TBL_Adressen
AdressID
Firma
Strasse
PLZ



TBL_PLZOrt
PLZ
Ort

Datenbanken

Normalisierung

N. Nazar
S. Baldes

Beispiel

Schlechtes Datenbank-Design

KundeNr	Name	PLZ	Stadt	AuftragNr	AuftragDatum	AuftragBeschreibung
1	Schulz GmbH	44701	Bochum	1	21.01.2005	Werkhalle streichen
2	Meier AG	47543	Bochum	2	22.01.2005	Teppichboden verlegen
1	Schulz GmbH	44701	Bochum	3	24.01.2005	Laminat verlegen
3	Fischer	44787	Bochum	4	27.01.2005	Tapezieren
1	Schulz GmbH	44702	Bochum	5	01.02.2005	Werkhalle Boden verlegen
2	Meier GmbH	47543	Bochum	6	01.02.2005	Laminat verlegen

Probleme:

- Die Datenbank enthält **redundante Daten**. Die Daten der Kunden mit mehreren offenen Aufträgen werden auch mehrfach abgespeichert. Dies verschwendet Speicherplatz.
- Die Datenbank enthält **widersprüchliche Daten**.

AuftragNr 3,5 → PLZ, Eingabefehler?

AuftragNr 2,6 → Rechtsform, Tippfehler / Rechtsform geändert?

Beispiel

Schlechtes Datenbank-Design


KundeNr	Name	PLZ	Stadt	AuftragNr	AuftragDatum	AuftragBeschreibung
1	Schulz GmbH	44701	Bochum	1	21.01.2005	Werkhalle streichen
2	Meier AG	47543	Bochum	2	22.01.2005	Teppichboden verlegen
1	Schulz GmbH	44701	Bochum	3	24.01.2005	Laminat verlegen
3	Fischer	44787	Bochum	4	27.01.2005	Tapezieren
1	Schulz GmbH	44702	Bochum	5	01.02.2005	Werkhalle Boden verlegen
2	Meier GmbH	47543	Bochum	6	01.02.2005	Laminat verlegen

Weitere, nicht sofort ersichtliche Problematiken:

- **Löschanomalie:**
Die Daten eines Kunden verschwinden, wenn der letzte Auftrag des Kunden gelöscht wird.
- **Einfügeanomalie:**
Einen neuen Kunden kann man nur dann eintragen, wenn er einen Auftrag erteilt.
- **Änderungsanomalie:**
Wenn man Daten eines Kunden ändern will, so muss man das so häufig machen, wie dieser Kunde Aufträge in der Datenbank hat.

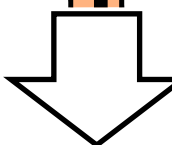
Beispiel

Besseres Datenbank-Design



KundeNr	Name	PLZ	Stadt	AuftragNr	AuftragDatum	AuftragBeschreibung
1	Schulz GmbH	44701	Bochum	1	21.01.2005	Werkhalle streichen
2	Meier AG	47543	Bochum	2	22.01.2005	Teppichboden verlegen
1	Schulz GmbH	44701	Bochum	3	24.01.2005	Laminat verlegen
3	Fischer	44787	Bochum	4	27.01.2005	Tapezieren
1	Schulz GmbH	44702	Bochum	5	01.02.2005	Werkhalle Boden verlegen
2	Meier GmbH	47543	Bochum	6	01.02.2005	Laminat verlegen

Tabelle Kunde
mit Primärschlüssel
KundeNr



KundeNr	Name	PLZ	Stadt
1	Schulz GmbH	44701	Bochum
2	Meier AG	47543	Bochum
3	Fischer	44787	Bochum

Tabelle Auftrag
mit Primärschlüssel
AuftragsNr und
Fremdschlüssel
KundenNr

AuftragNr	AuftragDatum	AuftragBeschreibung	KundeNr
1	21.01.2005	Werkhalle streichen	1
2	22.01.2005	Teppichboden verlegen	2
3	24.01.2005	Laminat verlegen	1
4	27.01.2005	Tapezieren	3
5	01.02.2005	Werkhalle Boden verlegen	1
6	01.02.2005	Laminat verlegen	2

Normalisierung

- Die Normalisierung bezweckt die **redundanz- und damit anomaliefreie** Speicherung von Informationen innerhalb der Tabellen.
- Dabei werden 3 verschiedenen Stufen unterschieden:
 - 1. Normalform
 - 2. Normalform
 - 3. Normalform

1. Normalform

Eine Tabelle befindet sich in der *ersten Normalform*, wenn **jedes Datenfeld atomar** ist.

- (↪ Datenfelder enthalten nur einen **einzelnen Eintrag**.
- ↪ Keine Wiederholungen von Attributen.)

Umformungen

1. || Album, Interpret ||



|| Album || Interpret ||

2. Not That Kind || Anastacia || {Not That Kind, I'm Outta Love, Cowboy & Kisses}



Not That Kind || Anastacia || Not That Kind

Not That Kind || Anastacia || I'm Outta Love

Not That Kind || Anastacia || Cowboy & Kisses c

1. Normalform

CD_Lieder

CD_ID	Album	Titelliste
4711	Anastacia - Not That Kind	{1. Not That Kind, 2. I'm Outta Love, 3. Cowboys & Kisses}
4712	Pink Floyd - Wish You Were Here	{1. Shine On You Crazy Diamond}

Nicht-atomar:
Album und Interpret
in einer Spalte

Nicht-atomar:
Mehrere Titel

CD_Lieder

	Albumtitel	Interpret	Titel	Track_ID
4711	Not That Kind	Anastacia	Not That Kind	1
4711	Not That Kind	Anastacia	I'm Outta Love	2
4711	Not That Kind	Anastacia	Cowboys & Kisses	3
4712	Wish You Were Here	Pink Floyd	Shine On You Crazy Diamond	1

In erster
Normalform

Alle Datenfelder sind atomar

Übung 1



Überführen Sie die folgende Tabelle in die erste Normalform.

Autor	Vorname	Titel	BuchNr	Verlag, Ort
Erhard	Heinz	Lachen, Erst sein ist alles	123, 234	rororo, Hamburg
Kuhlmann	Gregor	SQL, C++	345, 567	rororo, Hamburg
Meier	Helmut	SQL	345	rororo, Hamburg
Bilke	Petra	SQL, PHP	235, 789	knowware, Hornbäk

Übung 1

Lösungsvorschlag

<u>Autor</u>	Vorname	Titel	<u>BuchNR</u>	Ort	Verlag
Erhard	Heinz	Lachen	123	Hamburg	rororo
Erhard	Heinz	Erst sein ist alles	234	Hamburg	rororo
Kuhlmann	Gregor	SQL	345	Hamburg	rororo
Kuhlmann	Gregor	C++	567	Hamburg	rororo
Meier	Helmut	SQL	345	Hamburg	rororo
Bilke	Petra	SQL	235	Hornbäk	knowware
Bilke	Petra	PHP	789	Hornbäk	knowware

Die Tabelle ist in erster Normalform. Alle Datenfelder sind atomar.

Funktionale Abhängigkeit

Definition

Ein Attribut B ist von einem Attribut A *funktional abhängig*, wenn zu jedem Wert aus A **genau ein** Wert aus B existiert.

Wir schreiben dann

$$\mathbf{A \rightarrow B}$$
$$(\text{kenne } A \rightarrow \text{kenne } B)$$

Bemerkung:

Alle Attribute sind vom Gesamtschlüssel funktional abhängig.

Funktionale Abhängigkeit

Beispiel

<u>Kname</u>	Kfläche	<u>Lname</u>	Lfläche	Prozent
Europe	3234	Germany	357	100
Europe	3234	Russia	17075	20
Asia	44400	Russia	17075	80

Hier gelten die folgenden funktionalen Abhängigkeiten:

KName,LName → KFläche,LFläche,Prozent

Kname → KFläche

Lname → LFläche

Übung 2



Welche der folgenden funktionellen Abhängigkeiten werden durch die Tabelle erfüllt?

- **Reifentyp, Menge \rightarrow ReifenID**
- **Atnr \rightarrow Auftrag**
- **Atnr \rightarrow Reifentyp**
- **Menge \rightarrow Reifentyp**
- **Reifentyp \rightarrow Menge**
- **ReifenID \rightarrow Atnr**
- **ReifenID \rightarrow Auftrag**
- **ReifenID, Atnr \rightarrow Reifentyp, Menge, Auftrag**

ReifenID	Reifentyp	Menge	Atnr	Auftrag
1	Michelin	30	1	AuftragA
1	Michelin	30	2	AuftragB
1	Michelin	30	3	AuftragC
2	Pirelli	200	1	AuftragA
2	Pirelli	200	2	AuftragB
2	Pirelli	200	3	AuftragC
3	Dunlop	150	2	AuftragB
3	Dunlop	150	3	AuftragC
3	Dunlop	150	1	AuftragA
4	Good Year	300	3	AuftragC
5	Seat	300	1	AuftragA

Übung 2

Lösung

Welche der folgenden funktionellen Abhängigkeiten werden durch die Tabelle erfüllt?

- Reifentyp, Menge \rightarrow ReifenID
- Atnr \rightarrow Auftrag
- ~~Atnr \rightarrow Reifentyp~~
- ~~Menge \rightarrow Reifentyp~~
- Reifentyp \rightarrow Menge
- ~~ReifenID \rightarrow Atnr~~
- ~~ReifenID \rightarrow Auftrag~~
- ReifenID, Atnr \rightarrow Reifentyp, Menge, Auftrag

ReifenID	Reifentyp	Menge	Atnr	Auftrag
1	Michelin	30	1	AuftragA
1	Michelin	30	2	AuftragB
1	Michelin	30	3	AuftragC
2	Pirelli	200	1	AuftragA
2	Pirelli	200	2	AuftragB
2	Pirelli	200	3	AuftragC
3	Dunlop	150	2	AuftragB
3	Dunlop	150	3	AuftragC
3	Dunlop	150	1	AuftragA
4	Good Year	300	3	AuftragC
5	Seat	300	1	AuftragA

2. Normalform

Eine Tabelle befindet sich in der *zweiten Normalform*, wenn sie

- ↪ in 1. Normalform ist und
- ↪ kein Attribut ist funktional abhängig von einem Teilschlüssel. (**Teilschlüssel** \nrightarrow **Attribut**)

Alternative Formulierung:

Alle Attribute sind nur vom Gesamtschlüssel funktional abhängig.

Veranschaulichung:

Zusammenhängende Information wird in eine eigene Tabelle ausgelagert.

2. Normalform Beispiel

<u>Kname</u>	Kfläche	<u>Lname</u>	Lfläche	Prozent
Europe	3234	Germany	357	100
Europe	3234	Russia	17075	20
Asia	44400	Russia	17075	80

Kname → Kfläche

Lname → Lfläche

In zweiter
Normalform

<u>Kname</u>	Kfläche
Europe	3234
Asia	44400

<u>Kname</u>	<u>Lname</u>	Prozent
Europe	Germany	100
Europe	Russia	20
Asia	Russia	80

<u>Lname</u>	Lfläche
Germany	357
Russia	17075

Zusammenhängende Informationen werden in eigene Tabellen ausgelagert.

Übung 3



Der Primärschlüssel ist ein Kombinationsschlüssel aus Autor und ISBN.

Erfüllt die folgende Tabelle die zweite Normalform?

Falls nicht, erstellen Sie diese.

<u>Autor</u>	Vorname	Titel	<u>BuchNr</u>	Ort	Verlag
Erhard	Heinz	Lachen	123	Hamburg	rororo
Erhard	Heinz	Erst sein ist alles	234	Hamburg	rororo
Kuhlmann	Gregor	SQL	345	Hamburg	rororo
Kuhlmann	Gregor	C++	567	Hamburg	rororo
Meier	Helmut	SQL	345	Hamburg	rororo
Bilke	Petra	SQL	235	Hornbäk	knowware
Bilke	Petra	PHP	789	Hornbäk	knowware

Übung 3

Lösungsvorschlag

Die 2. Normalform ist nicht erfüllt, da der Titel nur von der ISBN abhängig ist und nicht von der Kombination aus Autor und ISBN.

<u>A_ID</u>	Autor	Vorname
1	Erhard	Heinz
2	Kuhlmann	Gregor
3	Meier	Helmut
4	Bilke	Petra

<u>BuchNr</u>	Ort	Verlag
123	Hamburg	rororo
234	Hamburg	rororo
345	Hamburg	rororo
567	Hamburg	rororo
235	Hornbäk	knowware
789	Hornbäk	knowware

<u>Buch_ID</u>	Titel	Autor	<i>BuchNr</i>
1	Lachen	1	123
2	Erst sein ist alles	1	234
3	SQL	2	345
4	C++	2	567
5	SQL	3	345
6	SQL	4	235
7	PHP	4	789

3. Normalform

Eine Tabelle befindet sich in der *dritten Normalform*, wenn sie

- ↪ in 2. Normalform ist und
- ↪ kein Nichtschlüssel-Attribut ist funktional abhängig von einem anderen. (**Attribut** \nrightarrow **Attribut**)

Veranschaulichung:

Kein Attribut kann den Wert eines anderen Attributes "voraussagen".

3. Normalform Beispiel

<u>SNr</u>	SName	LCode	Fläche
7	Freiburg	D	357
9	Berlin	D	357
40	Moscow	RU	17075
43	St.Petersburg	RU	17075

LCode → Fläche

In dritter
Normalform

<u>SNr</u>	SName	LCode
7	Freiburg	D
9	Berlin	D
40	Moscow	RU
43	St. Petersburg	RU

<u>LCode</u>	Fläche
D	357
RU	17075

Attribut ↗ Attribut

Übung 4



Prüfen Sie, ob die folgende Tabelle die dritte Normalform erfüllt?

Falls nicht, erstellen Sie diese.

<u>A_ID</u>	Autor	Vorname
1	Erhard	Heinz
2	Kuhlmann	Gregor
3	Meier	Helmut
4	Bilke	Petra

<u>BuchNr</u>	Ort	Verlag
123	Hamburg	rororo
234	Hamburg	rororo
345	Hamburg	rororo
567	Hamburg	rororo
235	Hornbäk	knowware
789	Hornbäk	knowware

<u>Buch_ID</u>	Titel	Autor	BuchNr
1	Lachen	1	123
2	Erst sein ist alles	1	234
3	SQL	2	345
4	C++	2	567
5	SQL	3	345
6	SQL	4	235
7	PHP	4	789

Übung 4

Lösungsvorschlag

Die Tabelle liegt nicht in der 3. Normalform vor, da der Ort vom Verlag abhängt.

<u>A_ID</u>	Autor	Vorname
1	Erhard	Heinz
2	Kuhlmann	Gregor
3	Meier	Helmut
4	Bilke	Petra

<u>Buch_ID</u>	Titel	Autor	BuchNr
1	Lachen	1	123
2	Erst sein ist alles	1	234
3	SQL	2	345
4	C++	2	567
5	SQL	3	345
6	SQL	4	235
7	PHP	4	789

<u>Verlag</u>	Ort	<u>BuchNr</u>	Verlag
rororo	Hamburg	123	rororo
knowware	Hornbäk	234	rororo
		345	rororo
		567	rororo
		235	knowware
		789	knowware

Übung 5



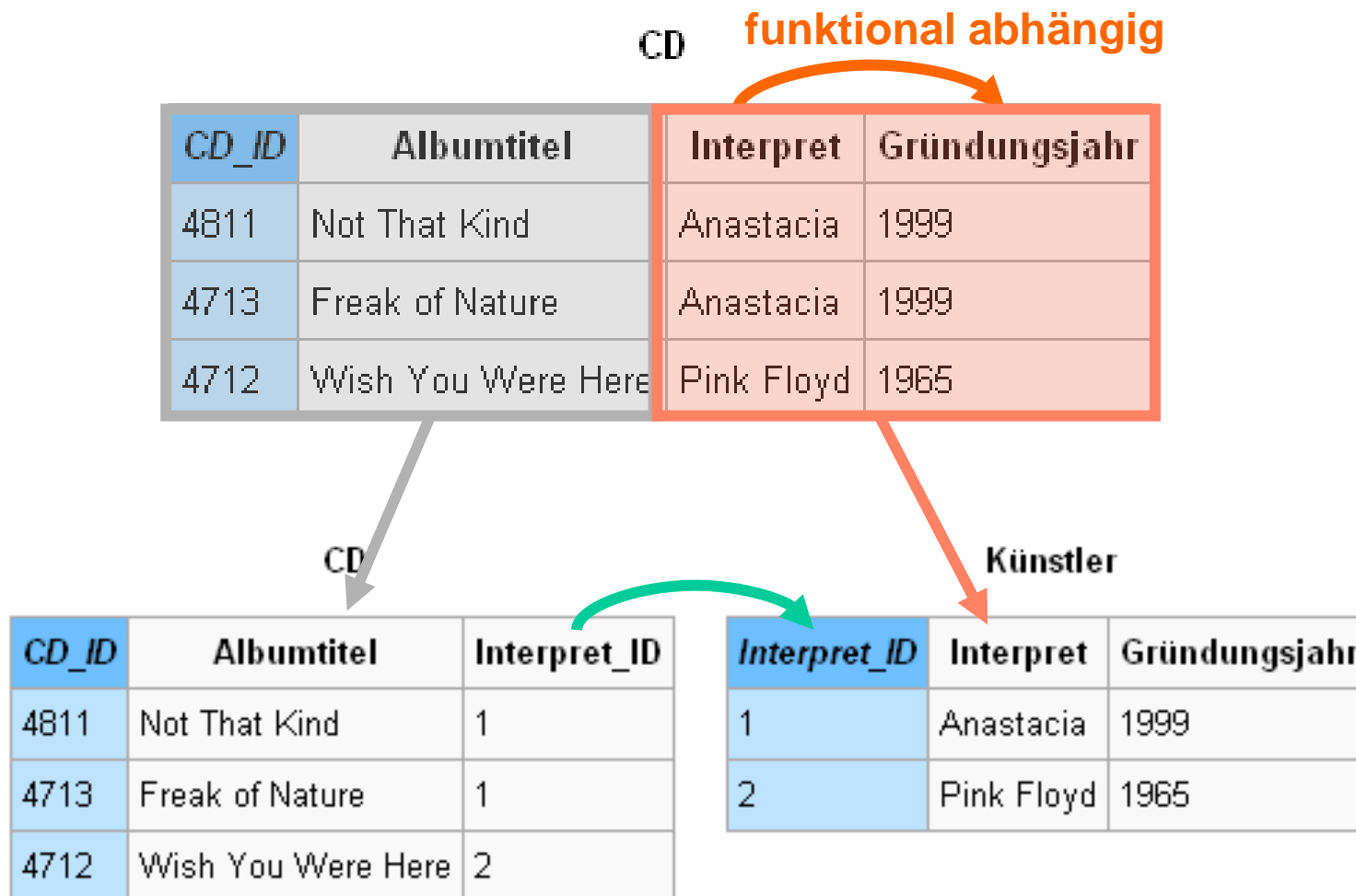
Überführen Sie die folgende Tabelle in die dritte Normalform. Bestimmen Sie dazu zuerst die funktionalen Abhängigkeiten.

CD

CD_ID	Albumtitel	Interpret	Gründungsjahr
4811	Not That Kind	Anastacia	1999
4713	Freak of Nature	Anastacia	1999
4712	Wish You Were Here	Pink Floyd	1965

Übung 5

Lösungsvorschlag



Übung 6

Komplex



Überführen Sie die folgende Tabelle in die dritte Normalform. Bestimmen Sie dazu zuerst die funktionalen Abhängigkeiten.

<u>kdnr</u>	name	vorname	vorwahl	tel	str	plz	ort	bland
1	Schumi	Michael	0761	123	S1	11	FR	BW
2	Alonso	Fernand	0761	543	S1	11	FR	BW
3	Massa	Felippe	0721	333	S2	12	KA	BW
4	Hamilton	Lewis	0721	456	S3	12	KA	BW
5	Lauda	Niki	089	874	S4	13	M	B
6	Heidfeld	Nick	089	744	S5	13	M	B
...
7	Franz	Müller	0721	325	S1	45	KA	BW
8	Heinz	Müller	0721	325	S1	45	KA	BW
9	Mike	Müller	0721	325	S1	45	KA	BW

Übung 6

Lösungsvorschlag

- Es gelten die folgenden funktionalen Abhängigkeiten:

Kdnr \rightarrow name, vname, vorwahl, tel, str, plz, ort, bland

plz \rightarrow vorwahl, ort, bland

vorwahl \rightarrow ort, bland

- Ergebnis der 3NF-Zerlegung:

Kunden(kdnr, name, vname, str, tel, plz)

Postleitzahl(plz, vorwahl)

Vorwahl(vorwahl, ort, bland)

Zusammenfassung

- **Normalisierung** verhindert Redundanz und damit Anomalien.
- 1NF: Alle Datenfelder sind **atomar**
- 2NF: **Teilschlüssel** \nrightarrow **Attribut**
- 3NF: **Attribut** \nrightarrow **Attribut**

Quellen

- Ahmad Nessar Nazar: Unterrichtsunterlagen
- Fabian Hauck: Unterrichtsunterlagen
- Stephan Baldes: Unterrichtsversuch Hr. Schlaich, 2007