



## Inhalt:

- *Einführung*
- *Normalformen 1NF bis 5NF*
- *Bedeutung von ERM*



*Die relationale Darstellung einer entworfenen Datenbank kann noch verfeinert werden. Die Basis für Verfeinerung bilden die Normalformen der Tabellen in der Datenbank.*

*Die Normalformen stellen eine Art der Qualität der Datenbank dar. Sie garantieren nur die notwendige Redundanz der Daten.*

*Es gibt Algorithmen, die eine nicht normalisierte Datenbank in eine normalisierte Form bringen. Mehrere Normalformen wurden im Laufe der Zeit entwickelt. Jede nächste Normalform basiert auf der vorigen.*

*Praktisch wichtig sind nur die ersten drei Normalformen.*

*Andererseits ist es manchmal sinnvoll, auf die Normalformen aus Performancegründen zu verzichten (z.B. lange Laufzeiten der Abfragen, große Anzahl der Tabellen).*



*Unter Normalisierung eines relationalen Datenbankmodells versteht man die Aufteilung von Attributen in mehrere Relationen (Tabellen) mithilfe der Normalisierungsregeln und deren Normalformen, sodass eine Struktur entsteht, die keine vermeidbaren Redundanzen mehr enthält.*

*Ziele der Datenbank-Normalisierung:*

- *Beseitigung von Redundanzen.*
- *Vermeidung von Anomalien (funktionelle und transitive Abhängigkeiten).*
- *Erstellung eines klar strukturierten Datenbankmodells.*



Die Grundlage der Normalisierung ist der Begriff funktionale Abhängigkeit oder entsprechend Primärschlüssel.

## Definition

Man betrachtet eine Relation,  $\alpha$  und  $\beta$  sind die Mengen der Attribute dieser Relation. Man sagt, dass  $\beta$  von  $\alpha$  funktional abhängig ist

$$\alpha \rightarrow \beta$$

wenn  $\forall$  Tupeln  $t_1$  und  $t_2$

mit  $t_1 \cap \alpha = t_2 \cap \alpha$  gilt auch  $t_1 \cap \beta = t_2 \cap \beta$

Die  $\beta$  hängt von der  $\alpha$  ab, wenn die  $\alpha$  die  $\beta$  eindeutig definiert.

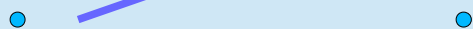
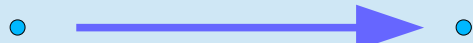
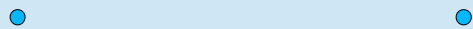
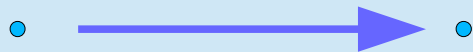
Man sagt, dass die  $\alpha$  alle anderen Attribute in der Relation bestimmt, d.h.  $\alpha$  ist ein Superschlüssel.

Ist die  $\alpha$  auch minimal, dann ist sie ein Kandidat für Primärschlüssel.



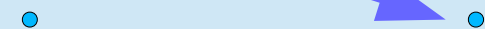
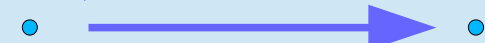
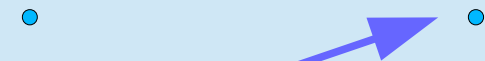
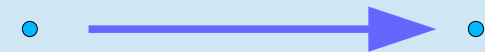
## Funktionale Abhängigkeit

$X \longrightarrow Y$



## Mehrwertige Abhängigkeit (multivalued)

$X \twoheadrightarrow Y$





Die Normalisierung ist eine Folge von Schritten zum Datenbankdesign, das einen effizienten Zugriff und Speicherung von Daten ermöglicht. Diese Schritte reduzieren die Datenredundanz und -inkonsistenz.

Normalformen:






- 1NF
- 2NF
- 3NF
- BCNF
- 4NF
- 5NF






Entzwei und gebietel! Tüchtig Wort;  
Verein und leite! Bessrer Hort. J. W. von Goethe


*divide et impera!*



## Definition ???

 <https://www.datenbanken-verstehen.de/datenmodellierung/normalisierung/erste-normalform/>  120%   Suchen

ken  Data Warehouse  Business Intelligence A-Z Lexikon  Videos  Übungen  Bücher

 » [Datenmodellierung](#) » [Normalisierung](#) » Erste Normalform (1NF)

## Erste Normalform (1NF)

Download Now

Play Now

Die **Erste Normalform (1NF)** ist dann gegeben, wenn **alle Informationen** in einer Tabelle **atomar vorliegen**.

Diesen Satz kann man in vielen [Datenbank Büchern](#) nachlesen, doch was bedeutet das wirklich?

Es bedeutet, dass **jede Information** innerhalb einer **Tabelle eine eigene Tabellenspalte bekommt** und **zusammenhängende Informationen**, wie zum Beispiel die Postleitzahl (PLZ) und der Ort, **nicht in einer Tabellenspalte vorliegen** dürfen.

A. Zimmermann

▪ HWR

▪ 2026

7/41



## Definition

*Eine Tabelle liegt in der ersten normalen Form (1NF) vor, wenn ihre Zellen nur atomare Werte beinhalten, d.h. sie enthalten nicht mehr als einen Wert (keine Auflistungen). "Atomar" bedeutet, dass die Werte nicht weiter in kleineren Komponenten zerlegt werden können, die einzeln einen Sinn im Anwendungsbereich ergeben.*

*Die 1NF vermeidet Wiederholungsgruppen, indem jede Gruppe in eine separate Tabelle eingefügt und durch eine 1:N-Beziehung verbunden wird.*

*Ist ein Attribut atomar oder nicht, ist stark von der Mini-Welt abhängig.*

*Man betrachte in zwei folgenden Beispielen immer dieselben Infos: PLZ, Straße, Hausnummer, Ort in zwei unterschiedlichen Mini-Welten.*





## Mini-Welt 1

*Sie arbeiten für eine Logistik-Firma (Müll-Abfuhr). Sie müssen die Routen für Ihre Autos optimal berechnen. Dafür müssen Sie die Adressen von Ihren Kunden zuerst nach PLZ sortieren, dann nach Straßen, danach nach Hausnummern. Nur so können Sie gewährleisten, dass Ihre Autos nicht zuerst in die PLZ 11222 fahren, dann in die PLZ 11333, und dann wieder nach PLZ 11222. Dasselbe betrifft auch die Straßen und Hausnummern.*

*Die PLZ, Straße und Hausnummer haben hier einen eigenen Sinn und Verwendung. Deswegen ist jede davon atomar.*



## Mini-Welt 2

*Sie arbeiten für eine Firma in der Buchhaltung und müssen monatlich die Gehaltsabrechnung an alle Mitarbeiter verschicken. Dafür brauchen Sie deren Adresse insgesamt. Die Hausnummer alleine, oder nur die Straße bringt Ihnen nichts. Wenn auf dem Briefumschlag nur die Straßen stehen, werden solche Briefe niemals zugestellt.*

*Hier hat nur die gesamte Adresse einen Sinn, deswegen ist sie nur insgesamt atomar.*



*Laut E.F.Codd müssen die Tabellen in dieser Form nicht unbedingt einen Primärschlüssel haben. Diese Anforderung ist nur für weitere Normalformen wichtig.*

*Eigentlich nachdem die Tabellen in 1NF gebracht sind, muss man sich Gedanken über Primärschlüssel machen.*



*Es sei eine Tabelle gegeben, um die Bahnwagen auf jeder Station zu überwachen.*

*WagenID: integer*

*z.B. 101*

*Beschreibung: varchar2(60)*

*z.B. 'Wagentyp, Leergewicht, Kapazität, Hersteller, Baujahr'*

*'Wg42, 3 t, 93 Personen, All4Rail, 2001'*

*Status: varchar2(16)*

*z.B. 'in Reparatur', 'im Test', 'im Betrieb'*

*Ankunft: datetime*

*z.B. '10:52'*

*Station: varchar2(20)*

*z.B. 'Alexanderplatz'*

*Primärschlüssel ist [WagenID,Ankunft].*

*Diese Tabelle ist nicht in 1NF, weil "Beschreibung" nicht atomar ist.*



*Das Feld [Beschreibung] wird in folgende Felder zerlegt:*

<i>WagenType</i>	<i>varchar(10)</i>
<i>Leergewicht</i>	<i>number</i>
<i>Kapazitaet</i>	<i>number</i>
<i>Hersteller</i>	<i>varchar(20)</i>
<i>Baujahr</i>	<i>datetime</i>

*Alle anderen Felder bleiben unverändert.*

*Jetzt befindet sich diese Tabelle in 1NF:*

*W1 = { [ WagenID: integer, Status: varchar2(16), Ankunft: datetime,  
Station: varchar2(20), WagenType: varchar(10),  
Leergewicht: number, Kapazitaet: number,  
Hersteller: varchar(20), Baujahr: datetime ] }*



Vertragsdatum	Kunde	Produkt
01.02.2010	Kohl	VW 30.000, BMW 40.000, Opel 40.000
03.01.2012	Schröder	VW 32.000, Mercedes 35.000
23.09.2013	Merkel	Toyota 25.000

	Vertragsdatum	Kunde	Produkt	Preis
◆	01.02.2010	Kohl	VW	30.000
★	03.01.2012	Schröder	VW	32.000
◆	23.09.2013	Merkel	Toyota	25.000
◆	01.02.2010	Kohl	BMW	40.000
◆	01.02.2010	Kohl	Opel	40.000
★	03.01.2012	Schröder	Mercedes	35.000



Vertragsdatum	Kunde	Produkt	Menge	Lieferadresse
01.02.2010	Schmidt	Stahlwalzen, Deutschland	17 Paletten	Wagendorf 20
03.01.2012	Stay & Go	Eis am Stiel, Italien	12 Kartons	Nordseeplatz 1
23.09.2013	Dirty GmbH	Seife, USA	2 Kartons	Bonner Platz 123
12.05.2013	Stay & Go	Handtuch, Pakistan	8 Paletten	Nordseeplatz 1
14.12.2012	Dirty GmbH	Handtuch, Indien	21 Paletten	Bonner Platz 123
15.06.2013	Zyperus KG	Gold, Südafrika	17 Barren	Zypern, Nikosia
15.06.2013	Schmidt	Gold, Namibien	12 Barren	Wagendorf 20



*Kein PK ist hier nötig*

Vertrags- datum	Kunde	Produkt	Herkunfts- land	V.art	Lieferadresse	Menge
01.02.2010	Schmidt	Stahlwalzen	Deutschland	Palette	Wagendorf 20	17
03.01.2012	Stay & Go	Eis am Stiel	Italien	Karton	Nordseeplatz 1	12
23.09.2013	Dirty GmbH	Seife	USA	Karton	Bonner Platz 123	2
12.05.2013	Stay & Go	Handtuch	Pakistan	Palette	Nordseeplatz 1	8
14.12.2012	Dirty GmbH	Handtuch	Indien	Palette	Bonner Platz 123	21
15.06.2013	Zyperus KG	Gold	Südafrika	Barren	Zypern, Nikosia	17
15.06.2013	Schmidt	Gold	Namibien	Barren	Wagendorf 20	12





### Definition

*Eine Tabelle liegt in der zweiten normalen Form (2NF) vor, wenn sie in der 1NF vorliegt und jedes Feld, das nicht zum Primärschlüssel gehört, hängt vom ganzen Primärschlüssel ab.*

*Die 2NF gewährleistet funktionale Abhängigkeit der Nichtschlüselfelder vom ganzen Schlüssel. Deswegen ist der Begriff "Primary Key" und eine Festlegung des Primärschlüssels absolut wichtig für die 2NF. Die Felder, die nur vom Teil des Primärschlüssels abhängig sind, werden in neuen Tabellen zusammen mit dem Teilschlüssel ausgelagert.*

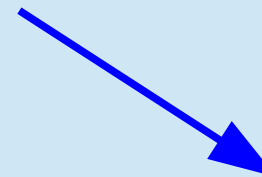
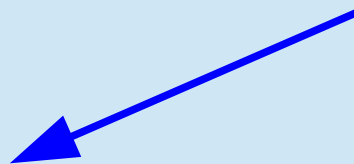
Regel: *Besteht der Primärschlüssel nur aus einem Feld, dann liegt die Tabelle in der 2NF automatisch vor.*



## Vorgang

PK = [S1, S2]

<u>S1</u>	<u>S2</u>		DepS2	



<u>S1</u>	<u>S2</u>		

<u>S2</u>	DepS2



Die Felder [Wagentyp], [Leergewicht], [Kapazität], [Hersteller], [Baujahr] hängen nur vom Feld [WagenID] ab, aber nicht vom Feld [Ankunft]. Deswegen befindet sich diese Tabelle nicht in 2NF.

Sie wird in 2NF gebracht, indem man sie entsprechend in zwei Tabellen mit unterschiedlichen PK aufteilt.

$W11 = \{[ \underline{WagenID}, \underline{Ankunft}, Status, Station ]\}$

$W12 = \{[ \underline{WagenID}, WagenType, Leergewicht, Kapazitaet, Hersteller, Baujahr ]\}$



PK wird festgelegt

<u>Vertrags- datum</u>	<u>Kunde</u>	<u>Produkt</u>	<u>Herkunfts- land</u>	V-Art	Lieferadresse	Menge
01.02.2010	Schmidt	Stahlwalzen	Deutschland	Palette	Wagendorf 20	17
03.01.2012	Stay & Go	Eis am Stiel	Italien	Karton	Nordseeplatz 1	12
23.09.2013	Dirty GmbH	Seife	USA	Karton	Bonner Platz 123	2
12.05.2013	Stay & Go	Handtuch	Pakistan	Palette	Nordseeplatz 1	8
14.12.2012	Dirty GmbH	Handtuch	Indien	Palette	Bonner Platz 123	21
15.06.2013	Zyperus KG	Gold	Südafrika	Barren	Zypern, Nikosia	17
15.06.2013	Schmidt	Gold	Namibien	Barren	Wagendorf 20	12



T-1

<u>Vertragsdatum</u>	<u>Kunde</u>	<u>Produkt</u>	<u>Herkunftsland</u>	Menge
01.02.2010	Schmidt	Stahlwalzen	Deutschland	17
03.01.2012	Stay & Go	Eis am Stiel	Italien	12
23.09.2013	Dirty GmbH	Seife	USA	2
12.05.2013	Stay & Go	Handtuch	Pakistan	8
14.12.2012	Dirty GmbH	Handtuch	Indien	21
15.06.2013	Zyprus KG	Gold	Südafrika	17
15.06.2013	Schmidt	Gold	Namibien	12



<u>Kunde</u>	Lieferadresse
Schmidt	Wagendorf 20
Stay & Go	Nordseeplatz 1
Dirty GmbH	Bonner Platz 123
Zyprus KG	Zypern, Nikosia

T-3

<u>Produkt</u>	<u>Herkunftsland</u>	V-Art
Stahlwalzen	Deutschland	Palette
Eis am Stiel	Italien	Karton
Seife	USA	Karton
Handtuch	Pakistan	Palette
Handtuch	Indien	Palette
Gold	Südafrika	Barren
Gold	Namibien	Barren

T-2

<u>Produkt</u>	<u>Herkunftsland</u>
Stahlwalzen	Deutschland
Eis am Stiel	Italien
Seife	USA
Handtuch	Pakistan
Handtuch	Indien
Gold	Südafrika
Gold	Namibien

Wenn [V-Art] nur  
von [Produkt]  
abhängig ist

T-3-1

T-3-2

<u>Produkt</u>	V-Art
Stahlwalzen	Palette
Eis am Stiel	Karton
Seife	Karton
Handtuch	Palette
Gold	Barren



RNr	ArtNr	LagerOrt	Anzahl
100100	1010	22	1
100100	1020	15	2
100100	1030	9	5
100103	1040	13	10
100104	1040	13	6

ArtNr : LagerOrt = 1 : 1

Mehrere Artikel  
von derselben ArtNr  
an einem LagerOrt

Schlüsselkandidaten: [RNr,ArtNr], [RNr,LagerOrt].

Die 2NF ist in beiden Fällen verletzt!

Ist PK = [RNr,ArtNr], so hängt [LagerOrt] vom Teilschlüssel [ArtNr] ab.

Ist PK = [RNr,LagerOrt], so hängt [ArtNr] vom Teilschlüssel [LagerOrt] ab.

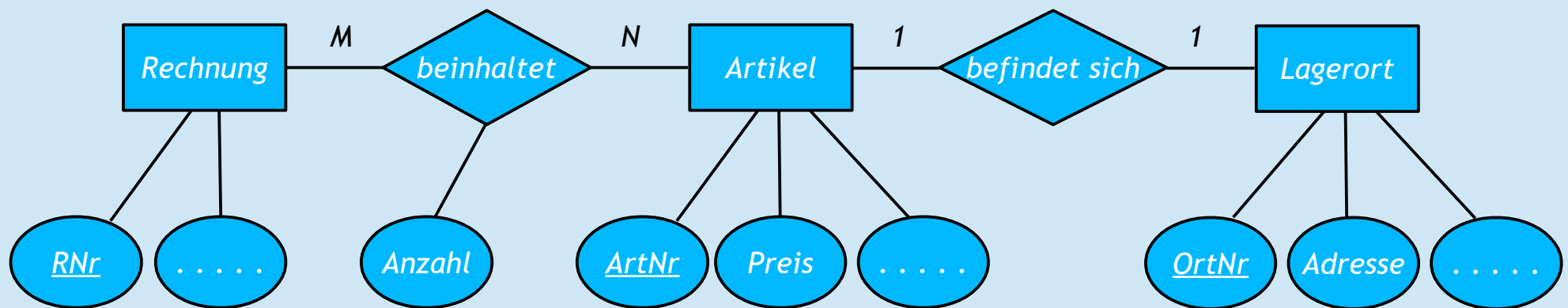
<u>RNr</u>	<u>ArtNr</u>	Anzahl
100100	1010	1
100100	1020	2
100100	1030	5
100103	1040	10
100104	1040	6

Tabelle ist konstant  
und  
hat weniger Zeilen  
als Ausgangs-Tabelle

<u>ArtNr</u>	<u>LagerOrt</u>
1010	22
1020	15
1030	9
1040	13
1050	5



Voriges Beispiel stammt eigentlich aus der unten durch ERM dargestellten Mini-Welt. Hier kommt man nie auf die vorige Tabelle, die die 2NF verletzt, sondern bekommt man sofort die Tabellen in der 3NF durch die Anwendung der Regeln. Ergebnis im vorigen Beispiel entspricht den beiden unteren Tabellen.



Rechnung	
<u>RNr</u>	...

Artikel		
<u>ArtNr</u>	Preis	...

Lagerort		
<u>OrtNr</u>	Adresse	...

beinhaltet		
<u>RNr</u>	<u>ArtNr</u>	Anzahl

befindet sich	
<u>ArtNr</u>	<u>OrtNr</u>





## Definition

Eine Tabelle liegt in der dritten normalen Form (3NF) vor, wenn sie in der 2NF vorliegt und alle Felder, die nicht zum Primärschlüssel gehören, sind von einander unabhängig.

Regel: Wenn lediglich ein Nichtschlüselfeld in der Tabelle vorhanden ist, dann liegt die Tabelle in der 3NF automatisch vor.

Die 3NF vermeidet die funktionale Abhängigkeiten unter den Feldern, die zu PK nicht gehören, durch Aufteilung der Tabelle.

In Praxis ist die 3NF oft ausreichend, um eine perfekte Balance aus Redundanz, Performance und Flexibilität für eine Datenbank zu gewährleisten.

Es gibt natürlich Sonderfälle, z.B. im wissenschaftlichen Bereich, wo eine Datenbank weiter normalisiert werden kann oder muss.

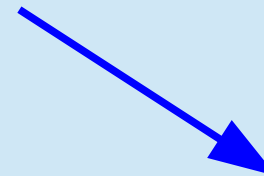
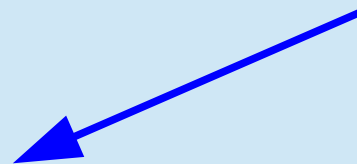
*The key,  
the whole key  
and nothing but the key.  
So help me Codd!*



## Vorgang

PK = [S1, S2]

<u>S1</u>	<u>S2</u>		F		DepF



<u>S1</u>	<u>S2</u>		F	

<u>F</u>	DepF



Die 2NF sah so aus:

Tabelle W11 ist schon in 3NF

W11 = {[ WagenID, Ankunft, Status, Station ]}

W12 = {[ WagenID, WagenType, Leergewicht, Kapazitaet, Hersteller, Baujahr ]}

In der zweiten Tabelle sind die Felder [Leergewicht], [Kapazitaet], [Hersteller] und u.U. auch [Baujahr] vom Feld [WagenType] abhängig, d.h. alle aufgelisteten Felder keine Bestandteile des Primärschlüssels sind und untereinander abhängig sind. Deswegen:

W121a = {[ WagenID, WagenType ]} [Baujahr] ist vom [WagenType] abhängig

W122a = {[ WagenType, Leergewicht, Kapazitaet, Hersteller, Baujahr ]}

oder

W121b = {[ WagenID, WagenType, Baujahr ]} [Baujahr] ist vom [WagenType] unabhängig

W122b = {[ WagenType, Leergewicht, Kapazitaet, Hersteller ]}



<u>KundenID</u>	Adresse	Telefon
117	Wespenplatz 2	121314
314	Storchstr. 20	212324
510	Bienenweg 18	313234

<u>KundenID</u>	Adresse
117	Wespenplatz 2
314	Storchstr. 20
510	Bienenweg 18

<u>KundenID</u>	Telefon
117	121314
314	212324
510	313234



## Definition

*Eine Tabelle liegt in der Boyce Codd normalen Form vor (BCNF), wenn sie in der 3NF vorliegt und kein Teil eines Schlüsselkandidaten funktional von keinem Teil eines anderen Schlüsselkandidaten abhängig ist.*

*Es geht hier nicht um Primärschlüssel alleine. BCNF behandelt die Schlüsselkandidaten, falls mehrere vorhanden sind, genau gesagt die Abhängigkeiten zwischen den Teilen dieser Schlüssel. Die Abhängigkeiten werden durch Aufteilung der Tabelle gelöst.*

*Die BCNF wendet man nur dann an, wenn mehrere Schlüsselkandidaten (inkl. Primärschlüssel) existieren und sich diese teilweise überlappen.*

Regel: *Gibt es in der Tabelle nur einen Schlüsselkandidat oder liegt keine Überlappung bei mehreren Schlüsselkandidaten vor, befindet sich die Tabelle automatisch in BCNF.*



<u>Name</u>	<u>Sportart</u> 1 : N	Verein
Schulz	Fußball	FC Berlin
Mayer	Fußball	FC Berlin
Zimmermann	Fußball	FC Marzahn
Mayer	Volleyball	VC Hamburg

Mini-Welt:  
 Ein Sportler kann  
 nur zu einem Verein  
 einer bestimmten Sportart gehören,  
 z.B. nur zu einem Fußballverein.  
 Aber zur gleichen Zeit kann er  
 auch zu einem Schachverein gehören  
 (das ist eine andere Sportart).

Eine Komponente des Primärschlüssels ist vom Nicht-Schlüssel-Feld abhängig: [Sportart] hängt vom [Verein] ab, nicht umgekehrt! Zu einer Sportart, z.B. Fußball, können mehrere Vereine gehören, aber ein Verein betreibt nur eine Sportart. Ursache – Schlüsselkandidat [Name, Verein].

<u>Name</u>	<u>Verein</u>
Schulz	FC Berlin
Mayer	FC Berlin
Mayer	VC Hamburg

Sportart	<u>Verein</u>
Fußball	FC Berlin
Volleyball	VC Hamburg



## Definition

Eine Tabelle liegt in der vierten normalen Form (4NF) vor, wenn sie in der BCNF vorliegt und nur die semantisch verbundene Nichtschlüsselattribute befinden sich in der Tabelle.

Die 4NF trennt die semantisch (thematisch, inhaltlich) unabhängigen Entitäten (Nichtschlüsselattribute) von einander durch Aufteilung der Tabelle.

In der 3NF werden die Felder [Leergewicht] und [Kapazitaet] viel öfter verwendet, als Felder [Hersteller] und [Baujahr], die eigentlich eher historische Bedeutung haben, deswegen werden sie höchstwahrscheinlich nie gleichzeitig verwendet.

$W122a1 = \{[ \underline{WagenType}, \text{Leergewicht}, \text{Kapazitaet} ]\}$

$W122a2 = \{[ \underline{WagenType}, \text{Hersteller}, \text{Baujahr} ]\}$

Unterschiedliche Semantik



### Definition

Eine Tabelle liegt in fünfter normalen Form (5NF) vor, wenn sie in der 4NF vorliegt und nicht mehr in Tabellen eines geringeren Grades zerlegt werden kann.

Die 5NF wird durch Aufteilung einer Tabelle erreicht. Die neuen Tabellen können jederzeit den ursprünglichen Zustand ohne Informationsverlust wieder herstellen (durch JOIN).

Die 5NF verlangt vereinfachte Tabellen, aus denen aber durch Projektions- und Verbundoperationen alle Informationen der ursprünglichen Tabelle wiederherstellbar sein müssen.

Nachteil:

Man braucht in Praxis ständig die ganzen Informationen, und man muss deswegen ständig die vereinfachten Tabellen wieder vereinigen.





Mehrere Mitarbeiter sind als Reiseführer für eine Touristik-Firma tätig. Sie können mehrere Touren durchführen. Diese Firma bestellt Busse und andere Verkehrsmittel für Reisen bei mehreren Transport-Unternehmen.

$T10 = \{[ \underline{PersID}, \underline{TourNr}, \underline{Trans-Unternehmen} ]\}$

In der Tabelle liegen keine mehrwertigen Abhängigkeiten vor, da die Felder [PersID], [TourNr] und [Trans-Unternehmen] zusammen eine informative Einheit bilden. Der primäre Schlüssel der Tabelle besteht aus allen drei Attributen.

Diese Tabelle kann man in die nächsten drei Tabellen zerlegen und wieder aus denen ohne Informationsverlust durch Operator Join herstellen.

$T11 = \{[ \underline{PersID}, \underline{TourNr} ]\}$

$T12 = \{[ \underline{PersID}, \underline{Trans-Unternehmen} ]\}$

$T13 = \{[ \underline{TourNr}, \underline{Trans-Unternehmen} ]\}$



<i>HerstID</i>	<i>ProduktID</i>	<i>PersID</i>
1	Stift	006
1	Ordner-L	007
1	Kopierpapier	007
2	Ordner-Z	006
3	Kopierpapier	007
1	Stift	003
3	Stift	006

$TA = \{[ \underline{HerstID}, \underline{ProduktID} ]\}$

$TB = \{[ \underline{PersID}, \underline{ProduktID} ]\}$

$TC = \{[ \underline{HerstID}, \underline{PersID} ]\}$



*Für Praxis ist es allgemein empfehlenswert, dass sich die Datenbank (genau gesagt – die Tabellen der Datenbank) in 3NF befinden soll. Die höheren Normalformen sind in theoretischen Untersuchungen wichtig.*

*Außerdem, in Praxis gelten viele Richtlinien in Unternehmen (z.B. über die Geschwindigkeit der Abfragen), die nur dann umgesetzt werden können, wenn man auf die 3NF (manchmal sogar auf die 2NF) verzichtet.*

*Wenn man von dem ERM ausgeht und den konzeptuellen Entwurf in den logischen laut den oben beschriebenen Regeln (inkl. Vereinfachung) umwandelt, dann ist es ziemlich wahrscheinlich, dass man die Tabellen in der 3NF bekommt. Klar, das kann man leider nicht 100%-tig garantieren.*

*Weiter wird gezeigt, wie die Normalisierung zu dem selben Ergebnis führt, wie auch die Konvertierung des ERM in die relationale Darstellung.*



*Es sei eine Tabelle gegeben, in der festgehalten wird, welche Mitarbeiter in welcher Abteilung arbeiten und an welchen Projekten sie beteiligt sind. Die Anforderungsanalyse und der konzeptuelle Entwurf wurden nicht so richtig und eingehend durchgeführt. Es ist aber aus den Kundeanforderungen bekannt, dass:*

- ein Mitarbeiter immer zu genau einer Abteilung gehört und kann an mehreren Projekten arbeiten;*
- eine Abteilung mehrere Mitarbeiter umfasst;*
- viele Mitarbeiter an einem Projekt tätig sind.*



Die folgende Tabelle kommt aus Praxis, also sie ist schon im Unternehmen vorhanden, z.B. als Excel-Tabelle:

<u>PersID</u>	Name	AbtNr	Abteilung	<u>ProjNr</u>	ProjBeschreibung
1	Anna	42	Second Level	BE	Bergland ...
1	Anna	42	Second Level	NO	Nordsee ...
2	Arnold	42	Second Level	NO	Nordsee ...
2	Arnold	42	Second Level	OG	Ostgipfel ...
3	Betty	53	First Dept	OG	Ostgipfel ...
3	Betty	53	First Dept	WO	West-Ozean ...
4	Chris	53	First Dept	WO	West-Ozean ...

Ist PK in Ordnung?

Welche NF ist es?

Sieht man hier die Redundanz?



*An dieser Tabelle sieht man bereits die Schwierigkeiten, die bei der Datenverarbeitung auftreten werden.*

- Einfügen. Soll ein neuer Mitarbeiter in eine bestehende Abteilung eingetragen werden, der bei einigen Projekten mitmachen soll, dann müssen mehrere Zeilen eingetragen werden, nämlich eine für jedes Projekt. Bei wenigen Operationen ist es kein Problem. Problematisch wird es beim Vertippen.
- Ändern. Soll ein Projekt umbenannt werden, müssen einige Zeilen geändert werden. Problematisch wird es beim Übersehen von Zeilen und beim Vertippen.
- Löschen. Soll ein Mitarbeiter gelöscht werden, müssen wieder mehrere Zeilen aus der Tabelle entfernt werden. Problematisch wird es beim Übersehen von Zeilen.



Die Normalisierung dieser Tabelle führt zu den Tabellen in 3NF:

<u>PersID</u>	Name	AbtNr
1	Anna	42
2	Arnold	42
3	Betty	53
4	Chris	53

<u>AbtNr</u>	Abteilung
42	Second Level
53	First Dept

<u>ProjNr</u>	ProjBeschreibung
BE	Bergland ...
NO	Nordsee ...
OG	Ostgipfel ...
WO	West-Ozean ...

<u>PersID</u>	<u>ProjNr</u>
1	BE
1	NO
2	NO
2	OG
3	OG
3	WO
4	WO



Wenn man aber mit ERM anfängt (also den konzeptuellen Entwurf richtig durchführt), dann bekommt man folgendes:

