# Datenbanken

**Teil 2: Informationen** 

**Kapitel 8: Normalformen** 



# Normalformenlehre zur Kontrolle der Tabellen einer Datenbank.

- 1. Redundanzen
- 2. Anomalien
- 3. Erste Normalform
- 4. Zweite Normalform
- 5. Dritte Normalform
- 6. Weitere Normalformen

**Definition: Redundanz** 

Wiederholung derselben Daten ohne tatsächlichen Informationsgewinn.

**Definition: Redundanz** 

Wiederholung derselben Daten ohne tatsächlichen Informationsgewinn.

# **Beispiel**

#### Schüler

ID_Schüler	Vorname	Nachname	ID_Klassenlehrer	Klassenlehrername
1	Markus	Schulte	1	Göbel
2	Willi	Kleinold	1	Göbel
3	Gabi	Müller	1	Göbel
4	Eva	Müller	2	Neitzel



**Definition: Redundanz** 

Wiederholung derselben Daten ohne tatsächlichen Informationsgewinn.

# **Beispiel**

#### Schüler

ID_Schüler	Vorname	Nachname	ID_Klassenlehrer	Klassenlehrername
1	Markus	Schulte	1	Göbel
2	Willi	Kleinold	1	Göbel
3	Gabi	Müller	1	Göbel
4	Eva	Müller	2	Neitzel
	•••		•••	

Hat man erst einmal dem Merkmal 'ID\_Klassenlehrer' mit Wert = '1' ...

**Definition: Redundanz** 

Wiederholung derselben Daten ohne tatsächlichen Informationsgewinn.

# **Beispiel**

#### Schüler

ID_Schüler	Vorname	Nachname	ID_Klassenlehrer	Klassenlehrername
1	Markus	Schulte	1	Göbel
2	Willi	Kleinold	1	Göbel
3	Gabi	Müller	1	Göbel
4	Eva	Müller	2	Neitzel
			•••	•••

Hat man erst einmal dem Merkmal 'ID\_Klassenlehrer' mit Wert = '1' den Klassenlehrernamen 'Göbel' zugeordnet,

• • •

**Definition: Redundanz** 

Wiederholung derselben Daten ohne tatsächlichen Informationsgewinn.

# **Beispiel**

#### Schüler

ID_Schüler	Vorname	Nachname	ID_Klassenlehrer	Klassenlehrername
1	Markus	Schulte	1	Göbel
2	Willi	Kleinold	1	Göbel
3	Gabi	Müller	1	Göbel
4	Eva	Müller	2	Neitzel
•••		•••		

Hat man erst einmal dem Merkmal 'ID\_Klassenlehrer' mit Wert = '1' den Klassenlehrernamen 'Göbel' zugeordnet, dann muss in jeder weiteren Zeile, in der der Wert '1' im Merkmal 'ID\_Klassenlehrer' erscheint,

**Definition: Redundanz** 

Wiederholung derselben Daten ohne tatsächlichen Informationsgewinn.

# **Beispiel**

#### Schüler

ID_Schüler	Vorname	Nachname	ID_Klassenlehrer	Klassenlehrername
1	Markus	Schulte	1	Göbel
2	Willi	Kleinold	1	Göbel
3	Gabi	Müller	1	Göbel
4	Eva	Müller	2	Neitzel
•••	•••			

Hat man erst einmal dem Merkmal 'ID\_Klassenlehrer' mit Wert = '1' den Klassenlehrernamen 'Göbel' zugeordnet, dann muss in jeder weiteren Zeile, in der der Wert '1' im Merkmal 'ID\_Klassenlehrer' erscheint, der Klassenlehrername wieder 'Göbel sein.

**Definition: Redundanz** 

Wiederholung derselben Daten ohne tatsächlichen Informationsgewinn.

# **Beispiel**

#### Schüler

ID_Schüler	Vorname	Nachname	ID_Klassenlehrer	Klassenlehrername
1	Markus	Schulte	1	Göbel
2	Willi	Kleinold	1	Göbel
3	Gabi	Müller	1	Göbel
4	Eva	Müller	2	Neitzel

Die wiederholte Nennung des Klassenlehrernamens 'Göbel' ist eine Redundanz OHNE tatsächlichen Informationsgewinn.

**Definition: Redundanz** 

Wiederholung derselben Daten ohne tatsächlichen Informationsgewinn.

# **Beispiel**

#### Schüler

ID_Schüler	Vorname	Nachname	ID_Klassenlehrer	Klassenlehrername
1	Markus	Schulte	1	Göbel
2	Willi	Kleinold	1	Göbel
3	Gabi	Müller	1	Göbel
4	Eva	Müller	2	Neitzel
•••	•••		•••	

Die wiederholte Nennung des Klassenlehrernamens 'Göbel' ist eine Redundanz OHNE tatsächlichen Informationsgewinn.

Redundanzen können zu Anomalien führen.

## Mutationsanomalie (Mutation der Daten):

Die Mutationsanomalie tritt auf, wenn die redundanten Daten (z.B. durch versehentliches Falschschreiben) verändert werden ("mutieren").

## Mutationsanomalie (Mutation der Daten):

Die Mutationsanomalie tritt auf, wenn die redundanten Daten (z.B. durch versehentliches Falschschreiben) verändert werden ("mutieren").

## Einfügeanomalie (ungewolltes Dateneinfügen):

Von einer Einfügeanomalie spricht man, wenn durch das Einfügen von Daten ungewollt weitere Daten eingefügt werden müssen.

### Mutationsanomalie (Mutation der Daten):

Die Mutationsanomalie tritt auf, wenn die redundanten Daten (z.B. durch versehentliches Falschschreiben) verändert werden ("mutieren").

## Einfügeanomalie (ungewolltes Dateneinfügen):

Von einer Einfügeanomalie spricht man, wenn durch das Einfügen von Daten ungewollt weitere Daten eingefügt werden müssen.

### Löschanomalie (unbeabsichtigtes Datenlöschen):

Eine Löschanomalie liegt vor, wenn durch das Löschen von Daten weitere Informationen verloren gehen, die aber gar nicht gelöscht werden sollten.

### Mutationsanomalie (Mutation der Daten):

Die Mutationsanomalie tritt auf, wenn die redundanten Daten (z.B. durch versehentliches Falschschreiben) verändert werden ("mutieren").

## Einfügeanomalie (ungewolltes Dateneinfügen):

Von einer Einfügeanomalie spricht man, wenn durch das Einfügen von Daten ungewollt weitere Daten eingefügt werden müssen.

### Löschanomalie (unbeabsichtigtes Datenlöschen):

Eine Löschanomalie liegt vor, wenn durch das Löschen von Daten weitere Informationen verloren gehen, die aber gar nicht gelöscht werden sollten.

## Änderungsanomalie (Datenänderung macht Folgeänderungen notwendig):

Wenn das Ändern eines Datensatzes zwangsläufig das Ändern weiterer Datensätze nach sich zieht, spricht man von einer Änderungsanomalie.

## Mutationsanomalie (Mutation der Daten):

Die Mutationsanomalie tritt auf, wenn die redundanten Daten (z.B. durch versehentliches Falschschreiben) verändert werden ("mutieren").

# **Beispiel**

#### Schüler

ID_Schüler	Vorname	Nachname	ID_Klassenlehrer	Klassenlehrername
1	Markus	Schulte	1	Göbel
2	Willi	Kleinold	1	Goebel
3	Gabi	Müller	1	Göbel
4	Eva	Müller	2	Neitzel
•••	•••	•••		

Der Name des Klassenlehrernamens "mutierte" auf Grund eines Tippfehlers von der richtigen Schreibweise 'Göbel' zu 'Goebel'.

### Einfügeanomalie (ungewolltes Dateneinfügen):

Von einer Einfügeanomalie spricht man, wenn durch das Einfügen von Daten ungewollt weitere Daten eingefügt werden müssen.

# **Beispiel**

#### Schüler

ID_Schüler	Vorname	Nachname	ID_Klassenlehrer	Klassenlehrername
1	Markus	Schulte	1	Göbel
2	Willi	Kleinold	1	Göbel
3	Gabi	Müller	1	Göbel
4	Eva	Müller	2	Neitzel
5	?	?	3	Neumann

Der Klassenlehrer 'Neumann' ist bereits bekannt und soll in die Datenbank eingetragen werden. Wenn der Klasse noch kein Schüler zugewiesen worden ist, wird ein Dummy-Schüler-Datensatz zum Eintragen des Klassenlehrers notwendig.

## Löschanomalie (unbeabsichtigtes Datenlöschen):

Eine Löschanomalie liegt vor, wenn durch das Löschen von Daten weitere Informationen verloren gehen, die aber gar nicht gelöscht werden sollten.

# **Beispiel**

#### Schüler

ID_Schüler	Vorname	Nachname	ID_Klassenlehrer	Klassenlehrername
1	Markus	Schulte	1	Göbel
2	Willi	Kleinold	1	Göbel
3	Gabi	Müller	1	Göbel
4	Eva	Müller	2	Neitzel

Löscht man den letzten Schüler aus der Klasse, in der der Lehrer 'Neitzel' Klassenlehrer ist. Dann geht die Information über diese Klassenlehrerschaft ebenfalls verloren.

## Änderungsanomalie (Datenänderung macht Folgeänderungen notwendig):

Wenn das Ändern eines Datensatzes zwangsläufig das Ändern weiterer Datensätze nach sich zieht, spricht man von einer Änderungsanomalie.

# **Beispiel**

#### Schüler

ID_Schüler	Vorname	Nachname	ID_Klassenlehrer	Klassenlehrername
1	Markus	Schulte	1	Müller
2	Willi	Kleinold	1	Müller
3	Gabi	Müller	1	Müller
4	Eva	Müller	2	Neitzel
•••				

Der Lehrer 'Göbel' lässt seinen Namen in 'Müller' ändern.

Diese eine Namensänderung führt zu sehr vielen Folgeänderungen.

... führen zu Anomalien!



... führen zu Anomalien!



# **Anomalien** ...

... führen zu inkonsistenten Daten!



## Inkonsistenzen

... führen zu Anomalien!



## **Anomalien** ...

... führen zu inkonsistenten Daten!



# Inkonsistenzen

... sind der Tod einer jeden Datenbank!

... führen zu Anomalien!



## Anomalien ...

führen zu inkonsistenten Daten!



## Inkonsistenzen

... sind der Tod einer jeden Datenbank!

Möchte man eine Datenbank "retten", muss man die Redundanzen vermeiden.

... führen zu Anomalien!



## Anomalien ...

... führen zu inkonsistenten Daten!





## Inkonsistenzen

... sind der Tod einer jeden Datenbank!

Möchte man eine Datenbank "retten", muss man die Redundanzen vermeiden.

Die Normalformen helfen beim Vermeiden und Beseitigen von Redundanzen aus den Tabellen.

Alle Tabellen	
Beliebige unnormalisierte Tabellen	

Alle Tabellen Beliebige unnormalisierte Tabellen					
1. Normalform					
Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)					

Beliebige unnormalisierte Tabellen

### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)



Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

#### 1.Normalform



## Definition: Erste Normalform (1NF)

Eine Tabelle befindet sich in der erster Normalform, falls die Wertebereiche der Merkmale atomar sind.

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

## **Tabelle NICHT in 1. Normalform**



## **Fachbelegung**

<u>ID</u>	<u>Schüler</u>	Name	ID_Fach	Fachbezeichnung
1		Klein	1	Mathematik
			2	Deutsch
2		Jung	2	Deutsch
			3	Englisch

Listen

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

## Tabelle in 1. Normalform



## **Fachbelegung**

ID_Schüler	ID_Fach	Name	Fachbezeichnung
1	1	Klein	Mathematik
1	2	Klein	Deutsch
2	2	Jung	Deutsch
2	3	Jung	Englisch

Alle Tabellen Beliebige unnormalisierte Tabellen					
1. Normalform Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)					

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

#### 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

#### 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel



Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

#### 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

#### **Zweite Normalform**



#### Definition: Zweite Normalform (2NF)

Eine Tabelle ist in zweiter Normalform, wenn sie die 1NF erfüllt und wenn <u>alle</u> Nichtschlüsselmerkmale vom zusammengesetzten Schlüssel voll funktional abhängig sind.

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

#### 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

## Tabelle NICHT in zweiter Normalform (erfüllt 1NF)



## **Fachbelegung**

ID_Schü	ler ID_Fach	chüler ID_Fach Name	Fachbezeichnung	Note
1	1	1 Klein	Mathematik	2
1	2	2 Klein	Deutsch	3
2	2	2 Jung	Deutsch	2
2	3	3 Jung	Englisch	4

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

#### 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

## Tabelle NICHT in zweiter Normalform (erfüllt 1NF)



Fachbe	leauna
<i>i</i> delibe	cgang

ID_Schüle	r ID_Fach	Name	Fachbezeichnung	Note
1	1	Klein	Mathematik	2
1	2	Klein	Deutsch	3
2	2	Jung	Deutsch	2
2	3	Jung	Englisch	4

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

#### 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

## Tabelle NICHT in zweiter Normalform (erfüllt 1NF)



Fachbe	leauna
<i>i</i> delibe	cgang

ID_Schüler	ID_Fach	Name	Fachbezeichnung	Note
1	1	Klein	Mathematik	2
1	2	Klein	Deutsch	3
2	2	Jung	Deutsch	2
2	3	Jung	Englisch	4

Beliebige unnormalisierte Tabellen

Eachhologung

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

#### 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel



racinelegalig					
<u>ID</u>	<u>Schüler</u>	ID_Fach	Name	Fachbezeichnung	Note
1		1	Klein	Mathematik	2
1		2	Klein	Deutsch	3
2		2	Jung	Deutsch	2
2		3	Jung	Englisch	4

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

#### 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel



rachbeiegung /					
ID_	<u>Schüler</u>	ID_Fach	Name	<b>Fachbezeichnung</b>	Note
1		1	Klein	Mathematik	2
1		2	Klein	Deutsch	3
2		2	Jung	Deutsch	2
2		3	Jung	Englisch	4

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

# 2. Normalform

Merkmale voll funktional abhängig

Das Merkmal Note ist vom zusammengesetzten Schlüssel ID\_Schüler und ID\_Fach abhängig!

Tabelle NICHT in zweiter Normalform (

# Fachbelegung

ID_Schüler	ID_Fach	Name	Fachbezeichnung	Note
1	1	Klein	Mathematik	2
1	2	Klein	Deutsch	3
2	2	Jung	Deutsch	2
2	3	Jung	Englisch	4

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

# 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

# **Tabellen NICHT in zweiter Normalform (erfüllt 1NF)**



#### Schüler

ID_Schüler	Name
1	Klein
2	Jung

#### Fach-Note

ID_Schüler	ID_Fach	Fachbezeichnung	Note
1	1	Mathematik	2
1	2	Deutsch	3
2	2	Deutsch	2
2	3	Englisch	4

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

# 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

# **Tabellen in zweiter Normalform**



#### Schüler

ID_Schüler	Name
1	Klein
2	Jung

#### Fach

ID_Fach	Fachbezeichnung	
1	Mathematik	
2	Deutsch	
3	Englisch	

#### Note

ID_Sc	hüler II	D_Fach	Note
1	1		2
1	2		3
2	2		2
2	3	,	4

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

#### 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

# 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

#### 3. Normalform

Keine transitiven Abhängigkeiten

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

# 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

#### 3. Normalform

Keine transitiven Abhängigkeiten



Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

# 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

#### 3. Normalform

Keine transitiven Abhängigkeiten

# **Dritte Normalform**



Definition: Dritte Normalform (3NF)

Eine Tabelle ist in dritter Normalform, wenn sie die 2NF erfüllt und <u>kein</u> Nichtschlüsselmerkmal vom Schlüssel transitiv abhängig ist!

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

# 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

# 3. Normalform

Keine transitiven Abhängigkeiten

# **Tabelle NICHT in dritter Normalform (erfüllt 2NF)**



#### Schüler

ID_Schüler	Name	ID_Klassenlehrer	Klassenlehrername
1	Schulte	1	Müller
2	Kleinold	1	Müller
4	Müller	2	Neitzel

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

# 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

# 3. Normalform

Keine transitiven Abhängigkeiten





ID_Schüler	Name	ID_Klassenlehrer	Klassenlehrername
1	Schulte	1	Müller
2	Kleinold	1	Müller
4	Müller	2	Neitzel

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

# 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

# 3. Normalform

Keine transitiven Abhängigkeiten





Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

# 2. Normalform

Schüler

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

# 3. Normalform

Keine transitiven Abhängigkeiten



ID_Schüler	Name	ID_Klassenlehrer	Klassenlehrername
1	Schulte	1	Müller
2	Kleinold	1	Müller
4	Müller	2	Neitzel



Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

# 2. Normalform

Schüler

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

# 3. Normalform

Keine transitiven Abhängigkeiten



ID_Schüler	Name	ID_Klassenlehrer	Klassenlehrername
1	Schulte	1	Müller
2	Kleinold	1	Müller
4	Müller	2	Neitzel

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

#### 2. Normalform

Merkmale voll funktional abhängig vom zus Der Klassenlehrername

3. Normalform

Keine transitiven Abhängigkeiten

Der Klassenlehrername ist transitiv abhängig (über das Merkmal ID Klassenlehrer)

vom Merkmal

ID\_Schüler



ID_Schüler	Name	ID_Klassenlehrer	Klassenlehrername
1	Schulte	1	Müller
2	Kleinold	1	Müller
4	Müller	2	Neitzel



Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

# 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

# 3. Normalform

Keine transitiven Abhängigkeiten

# **Tabelle in dritter Normalform**



#### Lehrer

<b>ID_Klassenlehrer</b>	Klassenlehrername
1	Müller
2	Neitzel

#### Schüler

ID_Schüler	Name	ID_Klassenlehrer
1	Schulte	1
2	Kleinold	1
4	Müller	2

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

# 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

#### 3. Normalform

Keine transitiven Abhängigkeiten

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

#### 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

# 3. Normalform

Keine transitiven Abhängigkeiten

# **Boyce-Codd- Normalform**

Nur Abhängigkeiten vom Schlüssel zulassen

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

# 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

# 3. Normalform

Keine transitiven Abhängigkeiten

# **Boyce-Codd- Normalform**

Nur Abhängigkeiten vom Schlüssel zulassen

# 4. Normalform

keine Mehrwertabhängigkeit

Beliebige unnormalisierte Tabellen

#### 1. Normalform

Wertebereiche der Merkmale sind atomar (keine Mengen oder Listen)

# 2. Normalform

Merkmale voll funktional abhängig vom zusammenges. Schlüssel

# 3. Normalform

Keine transitiven Abhängigkeiten

# **Boyce-Codd- Normalform**

Nur Abhängigkeiten vom Schlüssel zulassen

# 4. Normalform

keine Mehrwertabhängigkeit

# 5. Normalform

Nur triviale Verbundabhängigkeiten