

M164_Datenbanken erstellen und Daten einfügen

Tag 1

RECAP m162

1. Wie werden Netzwerk-Beziehungen im logischen Modell abgebildet?

- Primärschlüssel (Eindeutige, Identifikation eines Datensatzes)
- Fremdschlüssel (Verknüpfung zwischen Tabellen, um Beziehungen herzustellen)
- M: N Beziehungen (Zwischentabelle, die zwei Tabellen miteinander verbindet)

2. Was sind Anomalien in einer Datenbasis? Welche Arten gibt es?

Sind unerwünschte Inkonsistenzen oder Probleme bei der Datenverwaltung in einer Datenbank. Entstehen durch unzureichende Normalisierung.

- Einfüge Anomalie (Neue Daten können nicht gespeichert werden, weil andere nicht vorhandene Daten notwendig sind-> Bestellung nicht erfasst, weil Kunden nicht existiert).
- Änderungsanomalie: Änderungen mehrfach durchgeführt werden und führen zu Inkonsistenzen. (Eine Preisänderung muss an vielen Stellen geändert werden, und ein Fehler kann unterschiedliche Preise verursachen)
- Löschanomalie: bei Löschen von einem, gehen ungewollte weitere Daten verloren.

3. Gibt es redundante "Daten"? Warum?

Redundante Daten = mehrfach gespeicherte Informationen. Sie entstehen oft durch schlechte Datenmodellierung.

Drei Datenmodelle

1. Konzeptionelles Datenmodell (ERM)
2. Logisches Datenmodell (ERD)
3. Physisches Datenmodell

ERM->

- dient der **fachlichen Modellierung** und ist unabhängig von der Implementierung
- es wird direkt aus den **Anforderungen abgeleitet** und beschreibt die relevanten Entitäten und deren Beziehungen.

ERD->

- Stellt eine Abbildung des ERM auf ein spezifisches Datenbanksystem dar.
- Ergänzt um technische Aspekte wie **Primär- und Fremdschlüssel**, sowie **Attribute**, jedoch **unabhängig** von einer bestimmten **Datenbanktechnologie**.

Physisches->

- **Erweitert** um technische Details, z.B. Datenbank spezifische Erzeugungsanweisungen (**DDL-Skripte**), **Indizes und Partitionierungen**.
- Dient als Grundlage für die tatsächliche **Implementierung** in der Datenbank

Normalisierung

Es existieren 5 Normalformen, die ersten drei werden angewandt.

Was ist Normalisierung?

Normalisierung ist ein Prozess in der Datenbankmodellierung, bei dem **Redundanzen reduziert** und die **Datenintegrität verbessert** werden. Ziel ist es, Daten effizient zu strukturieren, um **Inkonsistenzen und Anomalien** bei Einfüge-, Änderungs- und Löschoperationen zu **vermeiden**.

Folgend ein Bild mit Redundanzen (unsortiert)

Name	Anschrift	Wohnort	Rechnungsnummer und - Datum	Artikel	Rechnungsbetrag
Gustav Meier-Senn	Im Tobel 13	9248 Bichwil	Nr 001 01.02.2004	Socken, T-Shirt, T-Shirt	SFr 358.00
Ida Kuhler	Walchenstr. 45b	5224 Unterbözing	Nr 002 01.02.2004	Hosen, Hosen	SFr 90.00
Gustav Meier-Senn	Im Tobel 13	9248 Bichwil	Nr 004 20.03.2004	Hosen, Socken, Schuhe	SFr 750.00
Sylvia Brendel	Rennweg 7	8000 Zürich			
Gustav Meier-Senn	Im Tobel 13	9248 Bichwil	Nr 007 21.03.2004	T-Shirt, Jacke	SFr 320.00
Morani	Birkenweg 16	1265 La Cure	Nr 008 22.03.2004	Jacke, T-Shirt	SFr 40.00
Sylvia Benz	Im Tobel 13	5600 Baden			
Walter Morani	Birkenweg 16	1265 La Cure	Nr 009 22.03.2004	Pullover, Socken	SFr 42.00
Merz, Beat	Seehalde 34	5600 Lenzburg	Nr 010 25.03.2004	Brille, Brille, Pullover	SFr 520.00

Normalform 1

Erfüllt, wenn: Alle Attribute enthalten nur **atomare (unteilbare) Werte**.
Wiederholende Gruppen oder **mehrwertige Attribute** werden **vermieden**.

Abfolge:

- Alle zusammengesetzten Attribute wurden in eigene Attribute getrennt (Name-> Vorname und Nachname). Dies führt zu mehr Spalten als Ursprung.
- Alle Attribute, die mehrere Werte des gleichen Typs haben, wurden auf mehrere Zeilen getrennt (Artikel). Dies führt zu mehr Datensätzen.

Vorname	Nachname	Strasse	Haus-Nr./Plz	Wohnort	Rech.Nr	Datum	Artikel	Währung	Rechnungsbetrag
Gustav	Meier-Senn	Im Tobel	13/9248	Bichwil	001	01.02.2004	Socken	SFr	358
Gustav	Meier-Senn	Im Tobel	13/9248	Bichwil	001	01.02.2004	T-Shirt	SFr	358
Gustav	Meier-Senn	Im Tobel	13/9248	Bichwil	001	01.02.2004	T-Shirt	SFr	358
Ida	Kuhler	Walchenstr.	45b/5224	Unterbözing	002	01.02.2004	Hosen	SFr	90
Ida	Kuhler	Walchenstr.	45b/5224	Unterbözing	002	01.02.2004	Hosen	SFr	90
Gustav	Meier-Senn	Im Tobel	13/9248	Bichwil	004	20.03.2004	Hosen	SFr	750
Gustav	Meier-Senn	Im Tobel	13/9248	Bichwil	004	20.03.2004	Socken	SFr	750
Gustav	Meier-Senn	Im Tobel	13/9248	Bichwil	004	20.03.2004	Schuhe	SFr	750
Sylvia	Brendel	Rennweg	7/8000	Zürich					
Gustav	Meier-Senn	Im Tobel	13/9248	Bichwil	007	21.03.2004	T-Shirt	SFr	320
Gustav	Meier-Senn	Im Tobel	13/9248	Bichwil	007	21.03.2004	Jacke	SFr	320
	Morani	Birkenweg	16/1265	La Cure	008	22.03.2004	Jacke	SFr	40
	Morani	Birkenweg	16/1265	La Cure	008	22.03.2004	T-Shirt	SFr	40
Sylvia	Benz	Im Tobel	13/5600	Baden					
Walter	Morani	Birkenweg	16/1265	La Cure	009	22.03.2004	Pullover	SFr	42
Walter	Morani	Birkenweg	16/1265	La Cure	009	22.03.2004	Socken	SFr	42
Beat	Merz	Seehalde	34/5600	Lenzburg	010	25.03.2004	Brille	SFr	520
Beat	Merz	Seehalde	34/5600	Lenzburg	010	25.03.2004	Brille	SFr	520
Beat	Merz	Seehalde	34/5600	Lenzburg	010	25.03.2004	Pullover	SFr	520

Normalform 2

Erfüllt, wenn: Die erste Normalform abgeschlossen.
Alle **Nicht-Schlüsselattribute** hängen voll funktional vom **Primärschlüssel** ab (keine Teilabhängigkeiten).

Abfolge:

1. Entitäten identifizieren. Trennen sie mögliche Objekt-Typen voneinander. (zB Kunden, Rechnungen, Artikel)
2. Definiere sie für jede Entität einen Primärschlüssel. Sie dürfen gerne eine zusätzliche Spalte mit einer Laufnummer einfügen.
3. Kardinalitäten (Beziehungen) zwischen den Entitäten identifizieren.
4. Netzwerk Beziehungen auflösen und Foreign Key setzen.

Kunde						Rechnung		Artikel	Rechnung	
Vorname	Nachname	Strasse	Haus-Nr	PLZ	Wohnort	RechNr	Datum	Artikel	Währung	Rechnungsbetrag
Gustav	Meier-Senn	Im Tobel	13	9248	Bichwil	001	01.02.2004	Socken	SFr	358
Gustav	Meier-Senn	Im Tobel	13	9248	Bichwil	001	01.02.2004	T-Shirt	SFr	358
Gustav	Meier-Senn	Im Tobel	13	9248	Bichwil	001	01.02.2004	T-Shirt	SFr	358
Ida	Kuhler	Walchenstr.	45b	5224	Unterbözing	002	01.02.2004	Hosen	SFr	90
Ida	Kuhler	Walchenstr.	45b	5224	Unterbözing	002	01.02.2004	Hosen	SFr	90
Gustav	Meier-Senn	Im Tobel	13	9248	Bichwil	001	01.02.2004	Socken	SFr	358

Kunde							Rechnung			Artikel	
KundenId	Vorname	Nachname	Strasse	Haus-Nr	PLZ	Wohnort	RechnungId	Datum	Rechnungsbetrag	ArtikelId	Beschreibung
1	Gustav	Meier-Senn	Im Tobel	13	9248	Bichwil	1	01.02.2004	358	1	Snowboard
2	Ida	Kuhler	Walchenstr.	45b	5224	Unterbözing	2	01.02.2004	90	2	Schuh
3	Sylvia	Brendel	Rennweg	7	8000	Zürich	4	20.03.2004	750	3	Ski
4	Walter	Morani	Birkenweg	16	1265	La Cure	7	21.03.2004	320	4	Jacke
5	Beat	Merz	Seehalde	34	5600	Lenzburg	8	22.03.2004	40	5	Socken
							9	22.03.2004	42	6	Brille
							10	25.03.2004	520	7	Hosen
										8	T-Shirt
										9	Pullover
										10	Artikel ab
										11	Artikel cd
										12	Artikel ef
										13	Artikel gh

In Tabelle farbig markiert, die identifizierten Entitäten getrennt. Währung wurde weggelassen, da diese immer gleich ist. Wenn nicht gleich dann wieder einführen. Unten FKs eingefügt.

Kunde						Rechnung			Rechnungsartikel		Artikel	
Kundenid	Vorname	Nachname	Strasse	Haus-Nr	PLZ	Wohnort	Rechnungid	Datum	Rechnungsbetrag	Rechnungsartikelid	Artikelid	Beschreibung
1	Gustav	Meier-Senn	Im Tobel	13	9248	Bichwil	1	01.02.2004	358	1	1	Snowboard
2	Ida	Kuhler	Walchenstr.	45b	5224	Unterbözing	2	01.02.2004	90	2	2	Schuh
3	Sylvia	Brendel	Rennweg	7	8000	Zürich	4	20.03.2004	750	3	3	Ski
4	Walter	Morani	Birkenweg	16	1265	La Cure	7	21.03.2004	320	4	4	Jacke
5	Beat	Merz	Seehalde	34	5600	Lenzburg	8	22.03.2004	40	5	5	Socken
							9	22.03.2004	42	6	6	Brille
							10	25.03.2004	520	7	7	Hosen
										8	8	T-Shirt
										9	9	Pullover
										10	10	Artikel ab
										11	11	Artikel cd
										12	12	Artikel ef
										13	13	Artikel gh

Normalform 3

Erfüllt, wenn: Befindet sich in der zweiten Normalform und **kein Nichtschlüsselattribut transitiv** von einem Kandidatenschlüssel abhängig ist.

->bedeutet, dass jedes Attribut nur vom PK abhängen darf und nicht von anderen Attributen.

Beispiel: Wohnort ist vom PK-Kundengeld und auch von der PLZ abhängig. Deshalb wurde zusätzlich Spalte PLZ id hinzugefügt. Dies wurde so gemacht, denn einige PLZ mehrmals vorkommen für unterschiedliche Orte und sich dieser Werte daher nicht als PK eignet.

Kunde

KundenId	Vorname	Nachname	Strasse	Haus-Nr	PLZ	Wohnort
1	Gustav	Meier-Senn	Im Tobel	13	9248	Bichwil
2	Ida	Kuhler	Walchenstr.	45b	5224	Unterbözing
3	Sylvia	Brendel	Rennweg	7	8000	Zürich
4	Walter	Morani	Birkenweg	16	1265	La Cure
5	Beat	Merz	Seehalde	34	5600	Lenzburg

Kunde

KundenId	Vorname	Nachname	Strasse	Haus-Nr	PLZid
1	Gustav	Meier-Senn	Im Tobel	13	1
2	Ida	Kuhler	Walchenstr.	45b	2
3	Sylvia	Brendel	Rennweg	7	3
4	Walter	Morani	Birkenweg	16	4
5	Beat	Merz	Seehalde	34	5

Postcodes

PLZid	PLZ	Wohnort
1	9248	Bichwil
2	5224	Unterbözing
3	8000	Zürich
4	1265	La Cure
5	5600	Lenzburg
6	xxx	Weitere
7	xxx	Weitere