

Kapitel 3 - Normalisierung

- 3.1 Normalisierung
- 3.2 Erste, zweite und dritte Normalform
- 3.3 Anomalien
- 3.4 Zusatz: Vierte und fünfte Normalform

3.1 Normalisierung

3.1 Normalisierung

Definition

Normalisierung ist ein Verfahren zur Strukturierung von Daten in relationalen Datenbanken.

Ziel ist es, Redundanzen zu vermeiden und Daten konsistent zu speichern.

Ziele

Die Normalisierung soll:

- doppelte Daten vermeiden
 - Datenintegrität verbessern
 - Inkonsistenzen verhindern
 - Änderungsanomalien vermeiden
 - Tabellen sinnvoll aufteilen
-

Grundprinzip

Daten werden so auf mehrere Tabellen verteilt, dass jede Information möglichst nur einmal gespeichert wird.

Wichtige Normalformen

Normalform	Bedeutung
1. Normalform	atomare Werte
2. Normalform	volle Abhängigkeit vom gesamten Primärschlüssel
3. Normalform	keine Abhängigkeit von Nicht-Schlüsselattributen

Beispiel

Problem:

Kundendaten werden in jeder Bestellung erneut gespeichert.

Lösung:

Kundendaten werden in eine eigene Tabelle **Kunde** ausgelagert und über einen Fremdschlüssel mit **Bestellung** verbunden.

Merksatz

Normalisierung reduziert Redundanzen und verhindert widersprüchliche Daten.

Prüfungshinweis

In IHK-Aufgaben wird häufig gefragt, warum Tabellen aufgeteilt oder Fremdschlüssel verwendet werden.

Typische Fehler:

- Redundanzen nicht erkennen.
 - Normalisierung nur als Speicherplatzersparnis verstehen.
 - Fremdschlüssel nach dem Aufteilen vergessen.
 - N:M-Beziehungen nicht auflösen.
-

Quellen

- IT-Handbuch für Fachinformatiker

3.2 Erste, zweite und dritte Normalform

Definition

Normalformen beschreiben Regeln zur Strukturierung von Tabellen in relationalen Datenbanken.

Ziel ist es, Redundanzen und Anomalien zu vermeiden.

1. Normalform

Eine Tabelle befindet sich in der **1. Normalform**, wenn alle Werte atomar sind.

Das bedeutet:

- jedes Feld enthält nur einen Wert
- keine Aufzählungen in einem Feld
- keine mehrfachen Werte in einer Zelle

Beispiel falsch

Kundennr	Name	Telefonnummern
1	Müller	030123, 0176123

Beispiel richtig

Kundennr	Name	Telefonnummer
1	Müller	030123
1	Müller	0176123

2. Normalform

Eine Tabelle befindet sich in der **2. Normalform**, wenn sie in der 1. Normalform ist und jedes Nicht-Schlüsselattribut vom gesamten Primärschlüssel abhängig ist.

Wichtig bei:

- zusammengesetzten Primärschlüsseln
 - Zwischentabellen
 - N:M-Beziehungen
-

3. Normalform

Eine Tabelle befindet sich in der **3. Normalform**, wenn sie in der 2. Normalform ist und keine Nicht-Schlüsselattribute voneinander abhängig sind.

Das bedeutet:

- Attribute dürfen nur vom Schlüssel abhängen
- nicht von anderen Nicht-Schlüsselattributen

Übersicht

Normalform	Kerngedanke
1. Normalform	atomare Werte
2. Normalform	abhängig vom gesamten Primärschlüssel
3. Normalform	keine Abhängigkeit von Nicht-Schlüsselattributen

Merksatz

1NF = atomar

2NF = vom ganzen Schlüssel abhängig

3NF = nur vom Schlüssel abhängig

Prüfungshinweis

In IHK-Aufgaben werden meist die **1. bis 3. Normalform** abgefragt.

Typische Fehler:

- mehrere Werte in einem Feld zulassen
- zusammengesetzte Schlüssel nicht beachten
- Attribute in der falschen Tabelle lassen
- transitive Abhängigkeiten nicht erkennen

Hinweis

In der Datenbanktheorie gibt es weitere Normalformen, z. B. die **4. Normalform** und **5. Normalform**.

Für IHK-Aufgaben sind jedoch meistens die **1. bis 3. Normalform** entscheidend.

Quellen

- IT-Handbuch für Fachinformatiker

3.3 Anomalien

Definition

Anomalien sind Probleme, die durch ungünstige Tabellenstrukturen und redundante Daten entstehen können.

Sie treten häufig auf, wenn Daten nicht ausreichend normalisiert sind.

Arten von Anomalien

Anomalie	Bedeutung
Einfügeanomalie	Daten können nicht sinnvoll eingefügt werden
Änderungsanomalie	Änderungen müssen mehrfach durchgeführt werden
Löschanomalie	Beim Löschen gehen ungewollt weitere Informationen verloren

Einfügeanomalie

Eine **Einfügeanomalie** liegt vor, wenn neue Daten nicht gespeichert werden können, ohne andere Daten ebenfalls anzugeben.

Beispiel

Ein neuer Artikel kann nicht gespeichert werden, solange noch keine Bestellung dazu existiert.

Änderungsanomalie

Eine **Änderungsanomalie** liegt vor, wenn dieselbe Information an mehreren Stellen geändert werden muss.

Beispiel

Eine Kundenadresse ist mehrfach gespeichert und wird nur an einer Stelle geändert.

Löschanomalie

Eine **Löschanomalie** liegt vor, wenn beim Löschen eines Datensatzes auch wichtige andere Informationen verloren gehen.

Beispiel

Wird die letzte Bestellung eines Kunden gelöscht, verschwinden auch die Kundendaten.

Merksatz

Anomalien entstehen häufig durch Redundanz.

Prüfungshinweis

In IHK-Aufgaben wird häufig gefragt, warum Tabellen normalisiert oder aufgeteilt werden sollen.

Typische Fehler:

- Anomalien nur als technische Fehler verstehen.
 - Redundanz nicht als Ursache erkennen.
 - Löschanomalie und Änderungsanomalie verwechseln.
 - Normalisierung nicht als Lösung erkennen.
-

Quellen

- IT-Handbuch für Fachinformatiker

3.4 Zusatz: Vierte und fünfte Normalform

Definition

Neben der ersten, zweiten und dritten Normalform gibt es weitere Normalformen.

Für die IHK sind meistens die **1. bis 3. Normalform** wichtig.

Die **4. Normalform** und **5. Normalform** sind eher Zusatzwissen für komplexere Datenbankmodelle.

Übersicht

Normalform	Fachlicher Name	Kerngedanke
1. Normalform	1NF	atomare Werte
2. Normalform	2NF	abhängig vom gesamten Primärschlüssel
3. Normalform	3NF	keine Abhängigkeit von Nicht-Schlüsselattributen
BCNF	Boyce-Codd-Normalform	strengere Form der 3NF
4. Normalform	4NF	keine unabhängigen Mehrfachabhängigkeiten
5. Normalform	5NF / PJNF	keine verlustbehafteten Zerlegungen

4. Normalform

Eine Tabelle befindet sich in der **4. Normalform**, wenn sie in der 3. Normalform ist und keine unabhängigen Mehrfachabhängigkeiten enthält.

Eine Mehrfachabhängigkeit entsteht, wenn zu einem Datensatz mehrere voneinander unabhängige Werte gespeichert werden.

Beispiel

Ein Mitarbeiter kann mehrere Sprachen sprechen.

Ein Mitarbeiter kann mehrere Projekte bearbeiten.

Wenn Sprachen und Projekte unabhängig voneinander sind, sollten sie nicht gemeinsam in einer Tabelle gespeichert werden.

Besser

- Mitarbeiter
 - Mitarbeiter_Sprache
 - Mitarbeiter_Projekt
-

5. Normalform

Eine Tabelle befindet sich in der **5. Normalform**, wenn sie nicht weiter verlustfrei in kleinere Tabellen zerlegt werden kann.

Dabei geht es um sehr komplexe Beziehungen zwischen mehreren Tabellen.

Die 5. Normalform ist in der Praxis und in der IHK-Prüfung deutlich seltener relevant.

Merksatz

1NF bis 3NF sind für die IHK am wichtigsten.

4NF und 5NF sind Zusatzwissen für komplexere Datenbankstrukturen.

Prüfungshinweis

In IHK-Aufgaben werden normalerweise die **1. bis 3. Normalform** abgefragt.

Die 4. und 5. Normalform sollte man kennen, aber nicht so tief lernen wie die ersten drei Normalformen.

Typische Fehler:

- 4NF und 5NF für normale IHK-Aufgaben überlernen.
 - Mehrfachabhängigkeiten mit normalen Fremdschlüsseln verwechseln.
 - 5NF mit 3NF gleichsetzen.
 - Tabellen zu stark zerlegen, obwohl es für die Aufgabe nicht nötig ist.
-

Quellen

- IT-Handbuch für Fachinformatiker