

Normalisierung von Datenbanken



Normalisierung

Unter **Normalisierung** eines relationalen Datenschemas versteht man die Aufteilung von Attributen in mehrere Relationen gemäß den Normalisierungsregeln. Ziel ist ein Datenschema in einer Form, die

- Datenredundanzen vermeidet
- logische Widersprüche in der Datenbasis verhindert
- höchstmögliche Flexibilität und einen schnellen Zugriff gewährleistet

SNr	Name	Klasse	Klassenleiter	GTA_Nr	Beschreibung
1	Bucky Barnes	10b	Hr. Wilson	1,4	Schach, Fotografie
2	Steve Rogers	9a	Fr. Carter	2	Sport
3	Tony Stark	8c	Hr. Parker	1,3	Schach, Computer

Frei von Redundanzen bedeutet, dass sämtliche Daten **nur einmal erfasst** werden. Sprich man muss bei einer Änderung die Daten auch **nur einmal ändern**. Sollten redundante Informationen in einer Tabelle vorhanden sein, kann es zu **Anomalien** kommen. Grundsätzlich werden Anomalien in **3 verschiedene Arten** unterschieden:

①

Einfüge

-Anomalie: Datensatz kann nur eingefügt werden kann, wenn ein anderer Datensatz auch hinzugefügt wird. Z.B. eine GTA kann nur hinzugefügt werden, wenn auch ein jeweiliger Schüler hinzugefügt wird.

Lösch

-Anomalie: Datensätze, die aus Versehen mitgelöscht werden. So würden, wenn man den Schüler "Steve Rogers" löschen würde, alle Informationen über die GTA "Sport" ebenfalls verloren gehen.

Änderungs

-Anomalie: bei einer Änderung müssen immer mehrere Datensätze geändert werden. → Tippfehlern oder übersehen von Datensätzen. Wenn zum Beispiel die GTA "Schach" ihren Namen ändern würde, müsste dies in mehreren Zeilen geändert werden.

Normalformen

Um eine Datenbank nun **redundanzfrei** und **ohne Anomalien** zu gestalten, werden die **3 Normalformen** angewandt. Was diese sind und wann eine **Datenbank Normalisierung** fertig ist, folgt jetzt.

1. Normalform:

Eine Tabelle befindet sich in der ersten Normalform (1.NF), wenn sie nur elementare Attribute enthält und ein eindeutiger Primärschlüssel gegeben ist.



Prüfregeln:

R1: nur atomare Werte (eine Teilinformation pro Feld)
R2: Jeder Datensatz durch Primärschlüssel identifizierbar

Anwendung Regel 1

Problem: _____

Lösung: _____

Tabelle nach Anwendung der Regel:

SNr			Klasse	Klassenleiter	GTA_Nr	Beschreibung
1			10b	Hr. Wilson		
2			9a	Fr. Carter	2	Sport
3			8c	Hr. Parker		

Anwendung Regel 2

Problem: _____

Lösung: _____

2. Normalform:

Eine Relation ist in der zweiten Normalform (2.NF), wenn sie in der **1.NF** vorliegt und jedes Attribut, das nicht zum Primärschlüssel gehört, alle Attribute des Primärschlüssels zur eindeutigen Identifikation benötigt. D.h. **jedes Nichtschlüssel-Attribut muss voll funktional vom gesamten Primärschlüssel abhängig sein.**



Prüfregel:

R3: jedes Nichtschlüssel-Attribut muss voll funktional vom Schlüssel abhängig sein
→ d.h. **Attribut kann nur durch den gesamten Schlüssel eindeutig bestimmt werden**

Anwendung Regel 3

Problem: _____

Lösung: _____

Tabelle nach Anwendung der Regeln:

Tabelle:

Tabelle:

Tabelle: