



Normalisierung

Die wissenschaftliche Methode, einen Modellentwurf von Datenanomalien zu befreien, nennt man **Normalisierung**. In der Praxis bedeutet das das **schrittweise Zerlegen von Tabellen** (Relationen) auf der Grundlage funktionaler Abhängigkeiten. Die ersten drei Normalformen von E. F. Codd, dem Erfinder des relationalen Datenbankmodells, sind die gebräuchlichsten.

Erste Normalform (1NF)

Eine Relation (Tabelle) ist in der ersten Normalform, wenn sie **frei von Wiederholungsgruppen** ist **und** alle **Datenfelder** mit zusammengesetzten Inhalten in ihre einzelnen Bestandteile zerlegt, also **atomisiert** sind.

Setze diese Tabelle in die Erste Normalform.

| tblMeineCDs | | | |
|-------------|---|--------|---|
| CD-ID | Album | Heimat | Titelliste |
| 3572 | ZZ Top - Eliminator | Texas | 1. Gimme All Your Lovin'; 2. Got Me Under Pressure; 3. Sharp Dressed Man; 4. u.s.w. |
| 3124 | Hansi Hinterseer - Ein kleines Edelweiß | Alpen | 1. Wir kommen von den Bergen; 2. u.s.w. |
| 6332 | ZZ Top - Greatest Hits | Texas | 1. Gimme All Your Lovin'; 2. u.s.w. |

Analyse: Verletzung der 1NF

- Das Datenfeld *Album* beinhaltet die Attribute *Interpret* und *Albumtitel*.
- Das Datenfeld *Titelliste* enthält eine Menge von *Titeln*.

Probleme bei Abfragen:

- Das Datenfeld *Album* kann nicht nach dem *Albumtitel* sortiert werden.
- Die einzelnen Titel können nicht sortiert werden und nur als *Titelliste* angezeigt werden.

Lösungsvorschlag:

| tblMeineCDs | | | | | |
|-------------|----------------------|------------------|--------|----------|---------------------------|
| CD-ID | Albumtitel | Interpret | Heimat | Titel-Nr | Titel |
| 3572 | Eliminator | ZZ Top | Texas | 1 | Gimme All Your Lovin' |
| 3572 | Eliminator | ZZ Top | Texas | 2 | Got Me Under Pressure |
| 3572 | Eliminator | ZZ Top | Texas | 3 | Sharp Dressed Man |
| 3124 | Ein kleines Edelweiß | Hansi Hinterseer | Alpen | 1 | Wir kommen von den Bergen |
| 6332 | Greatest Hits | ZZ Top | Texas | 1 | Gimme All Your Lovin' |

Analyse: Tabelle ist in 1NF

- Das Attribut *Album* wurde in die Datenfelder *Albumtitel* und *Interpret* gespalten.
- Das Attribut *Titelliste* wurde in die Datenfelder *Titel-Nr* und *Titel* aufgeteilt.

In der überarbeiteten Tabelle gibt es weder Wiederholungsgruppen noch Datenfelder mit zusammengesetzten Attributwerten.