

Multidimensionales Datenmodell

Tobias Bittner

2024-12-11

- 1 Wiederholung
- 2 Hausaufgabe 1: Klassifikationshierarchie zur Dimension "Aufbauorganisation"
- 3 Hausaufgabe 2: Szenario "Data-Warehouse-Analyse von Einsatzdaten"
- 4 Hausaufgabe 3: ROLAP von Unternehmensverkaufszahlen
- 5 Hausaufgabe 4: Logisches Schema
- 6 Alte SQL-Klausuraufgaben

- 1 Wiederholung
- 2 Hausaufgabe 1: Klassifikationshierarchie zur Dimension "Aufbauorganisation"
- 3 Hausaufgabe 2: Szenario "Data-Warehouse-Analyse von Einsatzdaten"
- 4 Hausaufgabe 3: ROLAP von Unternehmensverkaufszahlen
- 5 Hausaufgabe 4: Logisches Schema
- 6 Alte SQL-Klausuraufgaben

■ Datenstrukturen

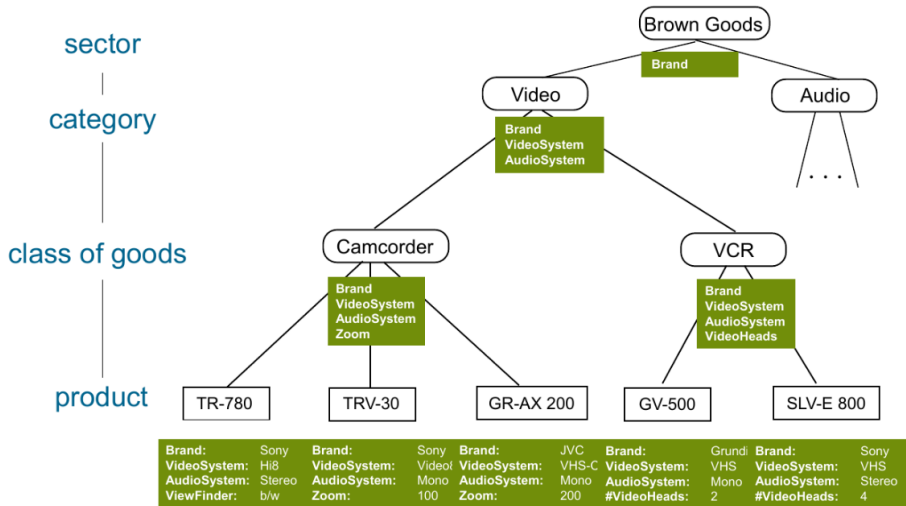
- **Qualifizierende Daten** – **Dimensionen** des Würfels
 - Klassenhierarchien mit qualifizierenden Attributen
 - Startpunkt für die Datenanalyse
- **Quantifizierende Daten** – **Zellen** des Datenwürfels
 - **Fakten** – z.B. Anzahl der Verkäufe, Umsatz etc.
 - Elementar-Bausteine für das Ergebnis der Datenanalyse

■ Ziele

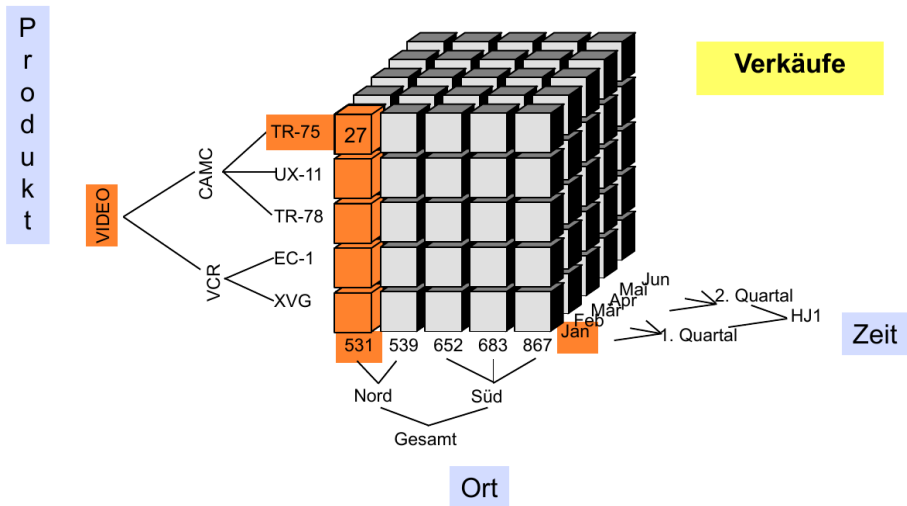
- Dimensionen müssen **unabhängig** sein
- Eindeutige Trennung von Fakten

Wiederholung - Dimensionen und Fakten II

Qualifizierende Daten – Klassenhierarchie mit Attributen

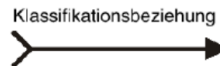
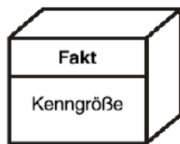


Wiederholung - Dimensionen und Fakten III



mE/R: Spezialisierung und Erweiterung des E/R-Modells

- Spezialisierung:
 - Neue Modellierungskonstrukte sind **Spezialfälle der bereits vorhandenen Konstrukte**
- Erweiterung
 - Neues Modell sollte **leicht erlernbar** sein
 - **Minimale Zahl** neuer Konstrukte
- Multidimensionale Semantik
 - Darstellung sollte **multidimensionale Semantik** der Daten umfassen



Fakten (Kenngrößen)

- können auch Eigenschaften zugesprochen werden.
- sind aber keine Datenstruktur an sich, eher analog zu einem Wertebereich (Beispiel: Maßeinheit bei biol./med./physik. Faktdaten)

Aggregationstyp

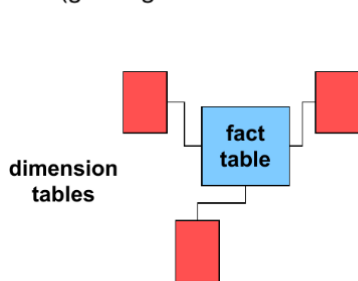
- nicht-triviale Eigenschaft neben Name und Wertebereich
 - definiert, welche Aggregationsoperationen auf einer Kenngröße ausgeführt werden dürfen!
- 1. beliebig aggregierbar (Sales, Turnover, ...): **FLOW**
- 2. nicht temporal summierbar (Stock, Inventory, ...): **STOCK**
- 3. nicht summierbar (Preis, Steuer, i.Allg. Faktoren): **VPU** (Value per Unit)
- (Immer durchgeführt werden dürfen MIN-, MAX-, und AVG-Operationen.)

Star Schema

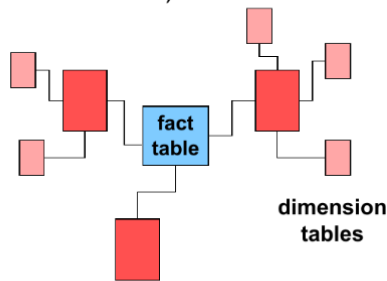
- Eine Tabelle für jede Dimension

Snowflake Schema

- Normalisierung der Dimensionstabellen
- Viele Tabellen je Dimension
(genaugenommen: eine Tabelle je Klassifikationsstufe)



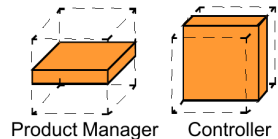
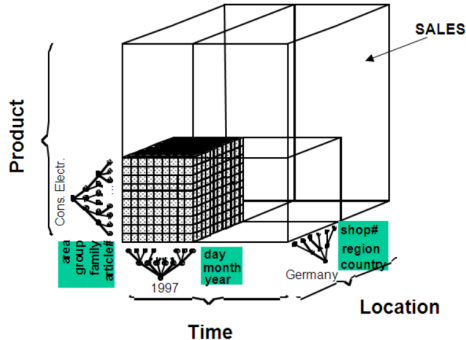
Star Schema



Snowflake Schema

■ Slicing und Dicing

- Selektion eines Teilwürfels: Dice
- Selektion von Scheiben aus einem Würfel: Slice

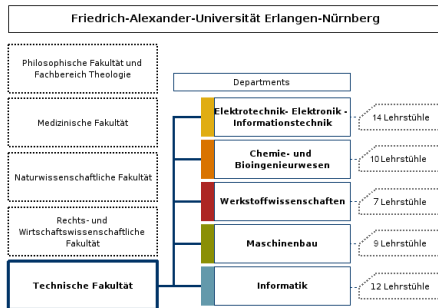


- 1 Wiederholung
- 2 Hausaufgabe 1: Klassifikationshierarchie zur Dimension "Aufbauorganisation"
- 3 Hausaufgabe 2: Szenario "Data-Warehouse-Analyse von Einsatzdaten"
- 4 Hausaufgabe 3: ROLAP von Unternehmensverkaufszahlen
- 5 Hausaufgabe 4: Logisches Schema
- 6 Alte SQL-Klausuraufgaben

Hausaufgabe 1: Klassifikationshierarchie zur Dimension "Aufbauorganisation" I

Erstellen einer Klassifikationshierarchie zu den Klassifikationsstufen:
A.Lehrstuhl → A.Department → A.Fakultät → A.Universität
Fokus auf der FAU und den Informatik Lehrstühlen der TechFak.
Gesucht: Relationen und Auszug aus Extension als:

- Star-Schema (einzelne Relation)
- Snowflake-Schema (normalisierte Tabellen)



- 1 Wiederholung
- 2 Hausaufgabe 1: Klassifikationshierarchie zur Dimension "Aufbauorganisation"
- 3 Hausaufgabe 2: Szenario "Data-Warehouse-Analyse von Einsatzdaten"**
- 4 Hausaufgabe 3: ROLAP von Unternehmensverkaufszahlen
- 5 Hausaufgabe 4: Logisches Schema
- 6 Alte SQL-Klausuraufgaben

Hausaufgabe 2: Szenario "Data-Warehouse-Analyse von Einsatzdaten" I

Gesucht: mE/R Entwurf für Pannendienst

Zusatzfrage: Aus wie vielen Attributen besteht der Primärschlüssel des Würfels bei der Umsetzung als Star/Snowflake-Schema?

- 1 Wiederholung
- 2 Hausaufgabe 1: Klassifikationshierarchie zur Dimension "Aufbauorganisation"
- 3 Hausaufgabe 2: Szenario "Data-Warehouse-Analyse von Einsatzdaten"
- 4 Hausaufgabe 3: ROLAP von Unternehmensverkaufszahlen
- 5 Hausaufgabe 4: Logisches Schema
- 6 Alte SQL-Klausuraufgaben

Hausaufgabe 3: ROLAP von Unternehmensverkaufszahlen I

Ein Unternehmen, das Elektronikgeräte verkauft, möchte seine Verkaufszahlen mittels ROLAP analysieren. Dazu sind nachfolgend ein multidimensionales Schema (logisches Modell) bzw. Star-Schema (physisches Modell) gegeben. Es soll untersucht werden, wie viele Artikel der Produktfamilie Waschgeräte im Jahr 2000 pro Monat in den unterschiedlichen Regionen Deutschlands verkauft wurden. Geben Sie das dazu erforderliche SQL-Statement an.

Hausaufgabe 3: ROLAP von Unternehmensverkaufszahlen II

Geographie-Tabelle

<u>GeographieID</u>
Filiale
Stadt
Region
Land

Produkt-Tabelle

<u>ProduktID</u>
Artikel
Familie
Gruppe
Branche

Absatz-Tabelle

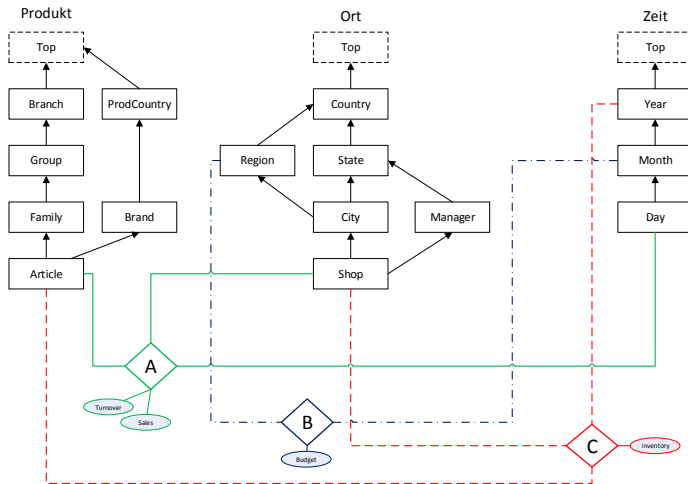
<u>ProduktID</u>
<u>GeographieID</u>
<u>ZeitID</u>
LAGERBESTAND
VERKAUEUFE
UMSATZ

Zeit-Tabelle

<u>ZeitID</u>
Tag
Monat
Jahr

- 1 Wiederholung
- 2 Hausaufgabe 1: Klassifikationshierarchie zur Dimension "Aufbauorganisation"
- 3 Hausaufgabe 2: Szenario "Data-Warehouse-Analyse von Einsatzdaten"
- 4 Hausaufgabe 3: ROLAP von Unternehmensverkaufszahlen
- 5 Hausaufgabe 4: Logisches Schema**
- 6 Alte SQL-Klausuraufgaben

Hausaufgabe 4: Logisches Schema I



Hausaufgabe 4: Logisches Schema II

- ① Wieviele Tabellen umfasst jeweils ein Snowflake-Schema für die Würfel A, B und C?
- ② Wieviele Tabellen umfasst ein Snowflake-Schema für den Würfel D ((Product.Brand, Ort.Region, Zeit.Jahr), (T,S))?
- ③ Wieviele Attribute hat jeweils die Faktentabelle in den Schemata zu A, B, C und D?
- ④ Wieviele Tabellen umfasst ein Snowflake-Schema für den Würfel E ((Produkt.Artikel, Ort.City AS Herstellungsort, Ort.Shop AS Verkaufsort), (Liefermenge))?
Beachten Sie, dass hier die gleiche Dimension in einem Würfel für unterschiedliche Zwecke mehrfach verwendet wird.

- 1 Wiederholung
- 2 Hausaufgabe 1: Klassifikationshierarchie zur Dimension "Aufbauorganisation"
- 3 Hausaufgabe 2: Szenario "Data-Warehouse-Analyse von Einsatzdaten"
- 4 Hausaufgabe 3: ROLAP von Unternehmensverkaufszahlen
- 5 Hausaufgabe 4: Logisches Schema
- 6 Alte SQL-Klausuraufgaben**