

1. Welche Möglichkeiten bestehen für die Umsetzung des mehrdimensionalen Datenwürfels im DWH? Welche Umsetzung ist besser?
2. Erläutern Sie die unterschiedlichen Summationstypen von Kennzahlen und geben Sie für beide Möglichkeiten geeignete abstrakte aber auch konkrete Beispiele an! Welche weiteren Eigenschaften müssen Kennzahlen erfüllen?
3. Überführen Sie einen anfragerelevanten Teil Ihres Parts des integrierten Datenbankschemas in ein relationales Data-Warehouse-Schema! Welche möglichen Umsetzungen bestehen?

Nehmen Sie dabei Bezug auf Dimensionen, die in ihrem lokalen Schema existieren, z. B. Zeit. Es soll versucht werden möglichst lange Drill-Pfade zu erzeugen.

Für die folgende Aufgabe sei anzunehmen, ihre Abteilung stelle das gesamte Data Warehouse dar.

- (a) Erstellen Sie ein Extrakt aus diesem Data Warehouse indem Sie eine materialisierte Sicht erzeugen. Dieser Extrakt kann z.B. die im Schema enthaltenen Daten für *ihre* lokale Finanzverwaltung o. Ä. aufbereiten.
Welche Vorteile und Notwendigkeiten ergeben sich durch die Verwendung materialisierter Sichten? Prüfen Sie ihre Thesen. Wird die von Ihnen erzeugte Sicht automatisch verwendet?
- (b) Führen Sie auf Basis dieser Daten im integrierten Data-Warehouse-Schema Anfragen aus, die einem Drill-Down und Roll-Up entsprechen.
- (c) Übertragen Sie die Vorgehensweise aus der Vorlesung zur Erstellung von Dimensionen auf das von Ihnen erstellte Schema.
- (d) Erstellen Sie auf Basis der Dimensionen die Anfrage zur Erzeugung einer Kreuztabelle und einer Detaillierung (Vergleiche Übungsblatt 3). Führen Sie auch auf Basis dieser Anfragen ein Drill-Down und Roll-Up durch. Was fällt ihnen auf?

ACHTUNG: Die Erfüllung dieser Aufgabe ist eine Voraussetzung für den Schein. Fügen Sie die Lösungen Ihrem bereits begonnenen SQL-Skript hinzu!

Tipp: Mittels PL/SQL können in Oracle sehr schnell Testdaten erzeugt werden.

4. Für die Modellierungsmöglichkeiten des Snowflake- und Star-Schema seien allgemeine Berechnungen bezüglich des Speicherbedarfs anzustellen:
 - (a) Gegeben seien zwei Dimensionen mit jeweils 7 Klassifikationsstufen – jede Stufe besitze ein Attribut. Die allgemeinste Klassifikationsstufe enthalte 2 Attribute. 4 Attributwerte einer untergeordneten Klassifikationsstufe sind einem Attributwert der entsprechend

PERSONAL_ID	NAME	Geselle	Lehrling
1	Mark	null	Verkaeuffer
2	Pete	Lagerist	null

Table 1: Zieldarstellung der daten

übergeordneten Klassifikationsstufe zugeordnet. Weiterhin sei eine Faktenmenge von 1000 Tupeln über ein Faktenattribut anzunehmen. Gehen sie davon aus, dass sämtliche Attribute ein byte groß sind und Schlüssel einelementig sind.

Berechnen Sie den Speicherbedarf für die Umsetzung als Star- und als Snowflake Schema.

- (b) Verändern Sie in der Berechnung des Speicherbedarfs den Parameter der möglichen Attributwerte auf 5 und ermitteln Sie den Speicherbedarf. Welche Tendenzen sind absehbar?

5. Aufgabe zur nächsten Übung:

In dieser Aufgabe soll das Wissen über Schema-Integration vertieft werden.

Führen Sie das folgende SQL-Skript in der Datenbank aus:

```
http://www.witi.cs.uni-magdeburg.de/iti\_db/lehre/dw/dwt1112/ueb04/ueb04\_prepare.sql
```

Überführen Sie die Daten der ersten beiden Tabellen (*Bierladen1*, *Bierladen2*) automatisiert in das Schema der dritten Tabelle *IntegratedBierladen*. Verlassen Sie sich nicht nur auf SQL, sondern schreiben Sie ein JAVA oder C++ Programm, welches die Transformation der Daten vornimmt.

Als Hilfestellung sei eine kleine Beispielanwendung "HelloJDBC.java" gegeben.¹ Für seine Anwendung muss der aktuelle Oracle-Treiber (ojdbc14-1.jar) in die Variable "CLASSPATH" aufgenommen werden.² Die *IntegratedBierladen*-Tabelle und die mit ihr verbundene Datendarstellung ist in Tabelle 1 dargestellt.

ACHTUNG: Die Erfüllung dieser Aufgabe ist eine Voraussetzung für den Schein.

¹http://www.witi.cs.uni-magdeburg.de/iti_db/lehre/dw/dwt1112/ueb04/ueb04_HelloJDBC.java

²http://www.witi.cs.uni-magdeburg.de/iti_db/lehre/dw/dwt1112/ueb04/ojdbc14-1.jar