
Aufgabe 1: [Referenzarchitektur]

a) Nennen Sie Gründe für den Einsatz eines Data-Warehouse-Systemes.

Zum Beispiel:

- *Analyse von Geschäftsprozessen soll laufenden Betrieb nicht stören.*
- *Analyse erfolgt über verteilte Datenquellen, deren Terminologie vereinheitlicht werden muss.*
- *Analyse erfolgt über historische Datenquellen.*
- *Analyse soll nachvollziehbar sein.*
- ...

b) Beschreiben Sie kurz die Funktionen eines Data Warehouse gemäß der Referenzarchitektur aus der Vorlesung.

Funktionen:

- *Unterstützung der Ladeprozesse*
Data Warehouse unterstützt die Ladewerkzeuge bei der Überführung der Daten aus der Basisdatenbank
- *Unterstützung der Analyseprozesse*
Data Warehouse unterstützt Anwendungen, wie z.B. OLAP- oder Data-Mining-Werkzeuge, bei der Analyse.

c) Vergleichen Sie Data Warehouse und Data Mart.

Data Warehouse ist nicht auf bestimmten Anwendungsbereich ausgerichtet, ein Data Mart hingegen schon. Ein Data Mart ist ein kleines Data Warehouse. Data Marts werden entweder von einem Data Warehouse abgeleitet oder stellen in ihrer Gesamtheit das Data Warehouse dar.

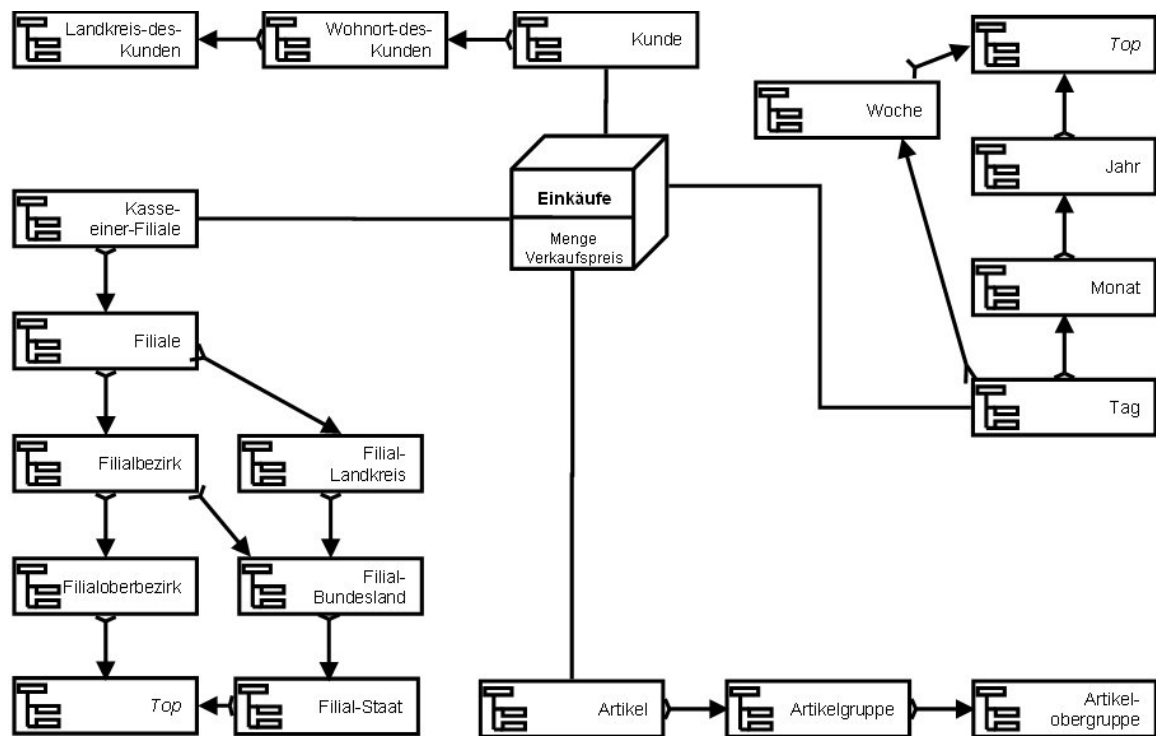
Aufgabe 2: [Data-Warehouse-Entwurf]

Sie sollen im Auftrag einer Sportartikelkette ein konzeptuelles Schema für ein Data Warehouse entwerfen. Dabei werden Einkäufe an den Kassen der einzelnen Filialen betrachtet.

Sie haben bereits folgende Dimensionsebenen identifiziert: *Kunde, Filialoberbezirk, Kasse-einer-Filiale, Landkreis-des-Kunden, Monat, Wohnort-des-Kunden, Filiale, Woche, Filial-Staat, Tag, Artikelobergruppe, Filial-Bundesland, Filialbezirk, Filial-Landkreis, Artikel, Jahr, Artikelgruppe*. Die Sportartikelkette organisiert ihre Filialen in Filialbezirken und diese wiederum in Filialoberbezirken. Bezirks- und Oberbezirksgrenzen haben nichts mit den politischen Grenzen in Deutschland zu tun. Sie wissen nur, dass kein Bezirk über die Grenzen eines Bundeslandes hinweg verläuft.

Als Fakten benötigen Sie nur *Menge* und *Verkaufspreis*.

a) Verwenden Sie eine der in der Vorlesung behandelten Modellierungssprachen für den konzeptuellen multidimensionalen Entwurf und entwerfen Sie ein geeignetes Schema. Modellieren Sie die vier Dimensionen *Zeit*, *Artikel*, *Kunde* und *Kasse/Filiale*.



b) Bestimmen Sie die Anzahl der verschiedenen Werte für alle Klassifikationsattribute der Zeit-Dimension. Betrachten Sie dabei nur das Jahr 2022.

Die Zeitdimension besteht aus den Kategorieattributen *Tag*, *Woche*, *Monat* und *Jahr*. Es gibt im Jahr 2007 365 Tage, 52 Wochen, 12 Monate und 1 Jahr. Es gibt also insgesamt $365+52+12+1 = 430$ verschiedene Werte.