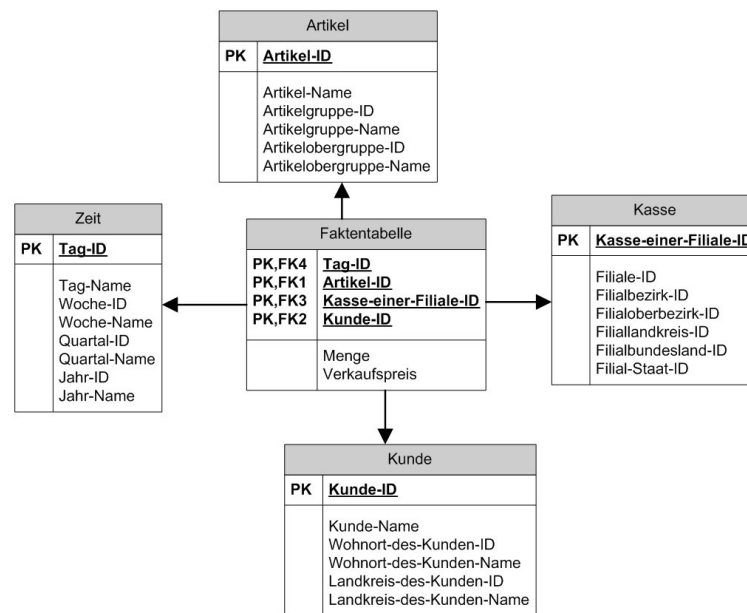
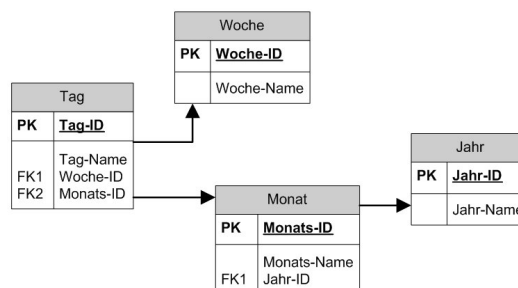


## Aufgabe 1: [Logische Modellierung]

a) Entwerfen Sie für das von Ihnen entworfene Modell aus Aufgabe 2 des 1. Übungsblattes ein Star-Schema.



b) Entwerfen Sie für das vorangegangene Modell einen Teil eines Snowflake-Schemas. Beschränken Sie sich dabei auf die Satellitentabellen der Zeit-Dimension.



c) Vergleichen Sie Vor- und Nachteile von Star- und Snowflake-Schema.

*Im Gegensatz zum Star-Schema sind im Snowflake-Schema Klassifikationsbeziehungen explizit modelliert und Dimensionstabellen normalisiert. Das Star-Schema hingegen ist hingegen übersichtlicher und im Allgemeinen performanter als das Snowflake-Schema.*

## Aufgabe 2: [Indexstrukturen]

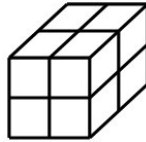


Abbildung 1: 2x2x2-Würfel

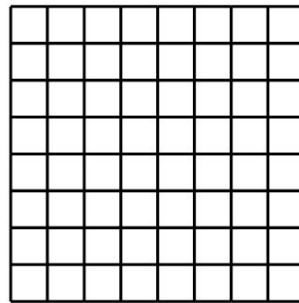
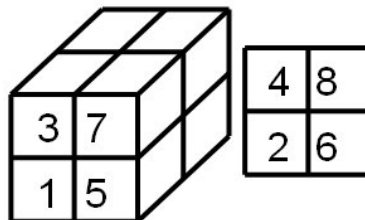


Abbildung 2: 8x8x1-Würfel

Verwenden Sie in dieser Aufgabe immer die gleiche Reihenfolge der Dimensionen:

1.  $\uparrow$ , 2.  $\nearrow$  (=“von vorne nach hinten“), 3.  $\rightarrow$

a) Weisen Sie anschließend jeder Würfelzelle aus Abbildung 1 und 2 eine eindeutige Nummer zu, indem Sie hierzu die Formel zur multidimensionalen Array-Speicherung verwenden.



8	16	24	32				
7	15	23	31				
6	14	22	30				
5	13	21	29				
4	12	20	28				
3	11	19	27				
2	10	18	26				
1	9	17	25				

b) Nennen Sie Gründe, warum man Indexstrukturen einsetzt.

*Indexstrukturen werden eingesetzt, um den Zugriff auf (im Allgemeinen sequentiell) gespeicherte Daten zu beschleunigen. In einer relationalen Datenbank vermeidet man damit die sequentielle Suche in allen Tupeln einer Tabelle.*



### Aufgabe 3: [DW-Materialisierung]

Nehmen Sie an, folgende Faktentabelle sei gegeben:

Monats-ID	Produkt-ID	Fahrzeug-ID	Einnahmen	Ausgaben
2004-Jan	5	1	1235	879
2004-Jan	7	3	5321	6345
2004-Feb	3	2	543	367
2004-Mar	4	1	235	198
2004-Apr	3	3	5432	5399
2004-Jan	2	2	745	4536
2004-Feb	2	3	346	636
2004-Jan	4	1	6246	3677
2004-Apr	1	2	326	436
2004-Apr	2	2	6436	7858
2004-May	4	3	8658	6356
2004-Jun	5	1	568	456
2004-May	4	3	5868	3167
2004-Jun	5	2	8762	6788

a) Was versteht man unter Materialisierung von Sichten?

*Unter Materialisierung versteht man das persistente Speichern von Sichten. Bei jedem Aufruf einer materialisierten Sicht wird die Sicht nicht jedes Mal neu berechnet, sondern die persistent gespeicherte Tabelle verwendet. Wird allerdings eine Basisrelation, auf die die Sicht basiert, geändert, so muss immer auch die materialisierte Sicht angepasst werden.*

b) Warum sollte man bestimmte Tabellen materialisieren?

*Das Durchsuchen der Faktentabelle und Berechnung von Aggregationen kostet Zeit. Wenn man häufig gebrauchte aggregierte Zwischenergebnisse materialisiert, d.h. physisch ablegt, speichert, so kann man dadurch einen erheblichen Performancegewinn erreichen.*

c) Was spricht dagegen? Warum werden nicht alle Tabellen materialisiert?

*Berücksichtigt man mehrstufige Aggregationshierarchien und mögliche Dimensionseinschränkungen, so gibt es exponentiell verschiedene Möglichkeiten Tabellen zu materialisieren. Der hierfür nötige Speicher und die auftretenden Änderungskosten rechtfertigen eine Materialisierung höchstens dann, wenn die Tabelle sehr häufig verwendet werden kann.*

d) Wie müsste eine materialisierte Sicht auf eine aggregierte Tabelle aussehen, wenn insbesondere der Gewinn auf Fahrzeugebene unabhängig von den transportierten Produkten häufig angefragt wird.

*Die Formulierung ist nicht eindeutig, deswegen gibt es zwei verschiedene Tabellen als Lösung. In beiden wird die Spalte Produkt ausgeblendet und Einnahmen und Ausgaben zu Gewinn verdichtet. In der ersten Tabelle bleiben Monats-ID und Fahrzeug-ID erhalten, in der zweiten Tabelle nur Fahrzeug-ID.*

Monats-ID	Fahrzeug-ID	Gewinn
2004-Jan	1	2925
2004-Jan	2	-3791
2004-Jan	3	-1024
2004-Feb	2	176
2004-Feb	3	-290
2004-Mar	1	37
2004-Apr	2	-1532
2004-Apr	3	33
2004-May	3	5003
2004-Jun	1	112
2004-Jun	2	1974

Abbildung 3: Lösungsvariante a)

Fahrzeug-ID	Gewinn
1	3074
2	-3173
3	3722

Abbildung 4: Lösungsvariante b)

### Aufgabe 4: [Kreuztabelle]

	Langensteinbach-Spedition	Müller-Ettingen	Schmidt-Bruchsal	Walter-Transport-Rheinstetten	DE123 Karlsruhe, Landkreis
Bekleidungszubehör.	0,00	0,00	215,84	0,00	215,84
Kopfbedeckungen.	0,00	0,00	211,17	0,00	211,17
Spezialkleidung.	382,54	0,00	388,10	0,00	770,64
Verschlüsse.	0,00	0,00	567,88	0,00	567,88
<b>Spezialkleidung und Zubehör.</b>	<b>382,54</b>	<b>0,00</b>	<b>1382,99</b>	<b>0,00</b>	<b>1765,53</b>

Abbildung 5: Kreuztabelle

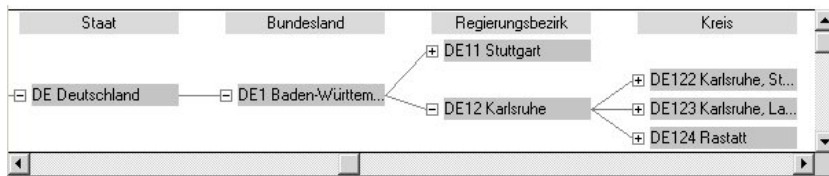


Abbildung 6: Auszug aus Ort-Dimension

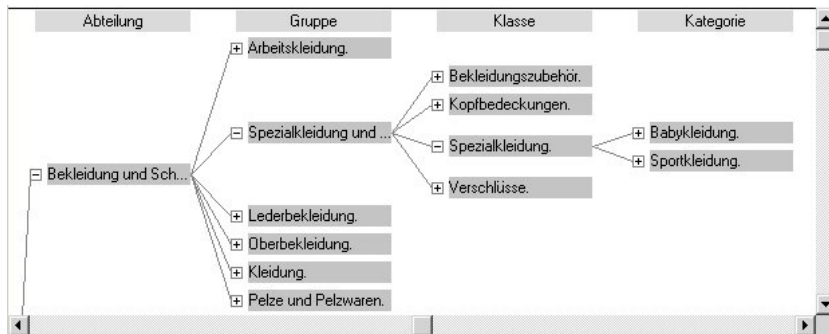


Abbildung 7: Auszug aus Produkt-Dimension

Wie lauten die Zeilen und Spalten der Kreuztabelle, wenn Sie jeweils folgende Operatoren ausführen?

a) Roll-up (“DE123 Karlsruhe, Landkreis”)

	DE122 Karlsruhe,	DE123 Karlsruhe, Landkreis	DE124 Rastatt	DE12 Karlsruhe
Bekleidungszubehör.	501,87	215,84	0,00	717,71
Kopfbedeckungen.	806,35	211,17	0,00	1017,52
Spezialkleidung.	153,71	770,64	0,00	924,35
Verschlüsse.	934,61	567,88	193,57	1696,06
<b>Spezialkleidung und Zubehör.</b>	<b>2396,54</b>	<b>1765,53</b>	<b>193,57</b>	<b>4355,64</b>

b) Slice(“Spezialkleidung”), Drill-down (“Spezialkleidung”)

	Langensteinbach- Spedition	Müller-Ettingen	Schmidt-Bruchsal	Walter-Transport- Rheinstetten	DE123 Karlsruhe, Landkreis
Babykleidung.	0,00	0,00	388,10	0,00	388,10
Sportkleidung.	382,54	0,00	0,00	0,00	382,54
<b>Spezialkleidung.</b>	<b>382,54</b>	<b>0,00</b>	<b>388,10</b>	<b>0,00</b>	<b>770,64</b>

c) Split (“Zeit”)

*Die hier gezeigte Kreuztabelle ist nur ein Ausschnitt der Lösung. Zusätzlich wurde die Zeitdimension auf das 3. Quartal 2003 beschränkt.*

	Langensteinbach-Spedition			
	Jul./2003	Aug./2003	Sep./2003	3. Q. 2003
Bekleidungszubehör.	0,00	0,00	0,00	0,00
Kopfbedeckungen.	0,00	0,00	0,00	0,00
Spezialkleidung.	0,00	382,54	0,00	382,54
Verschlüsse.	0,00	0,00	0,00	0,00
<b>Spezialkleidung und Zubehör.</b>	<b>0,00</b>	<b>382,54</b>	<b>0,00</b>	<b>382,54</b>

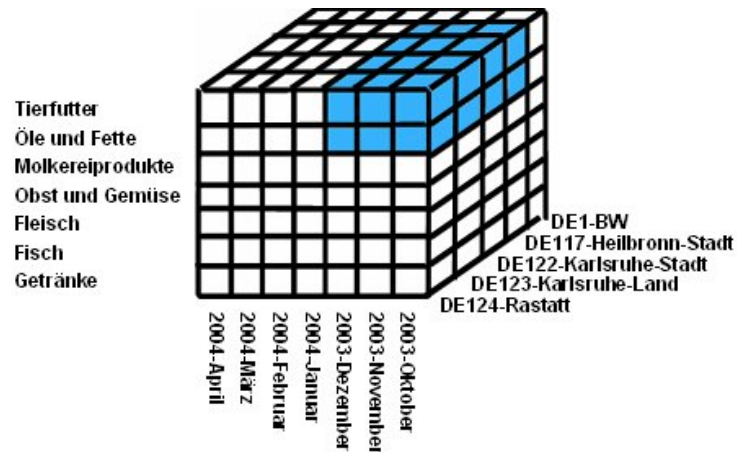


d) Split (“Zeit”), Merge (“Ort”)

*Für diese Kreuztabelle wurde die Zeitdimension zusätzlich auf das 3. Quartal 2003 beschränkt.*

	Juli/2003	Aug./2003	Sep./2003	3. Q. 2003
Bekleidungszubehör.	215,84	0,00	0,00	215,84
Kopfbedeckungen.	211,17	0,00	0,00	211,17
Spezialkleidung.	388,10	382,54	0,00	770,64
Verschlüsse.	173,54	0,00	394,34	567,88
<b>Spezialkleidung und Zubehör.</b>	<b>988,65</b>	<b>382,54</b>	<b>394,34</b>	<b>1765,53</b>

### Aufgabe 5: [Slice&Dice]

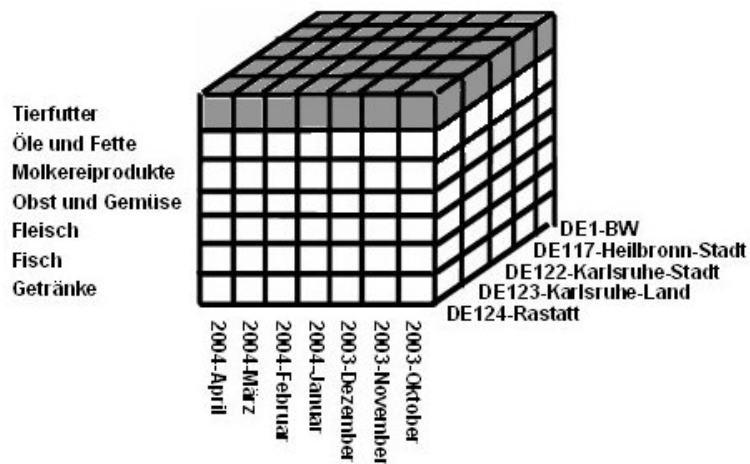


a) Erläutern Sie die Unterschiede zwischen den beiden OLAP-Operatoren *Slice* und *Dice*.

*Bei Slice wird der Ergebnisdatenraum einer Dimension auf einen Wert eingeschränkt, dadurch verringert sich die Dimensionalität. Bei Dice wird ebenfalls der Ergebnisdatenraum eingeschränkt, allerdings für eine oder mehrere Dimensionen und so, dass die Dimensionalität erhalten bleibt.*

b) Formulieren Sie in eigenen Worten eine *Slice*-Anweisung auf den gegebenen Würfel.

*Zeige mir die Fakten für Tierfutter für alle sieben Monate und alle Landkreise und Baden-Württemberg.*



c) Formulieren Sie in eigenen Worten eine *Dice*-Anweisung auf den gegebenen Würfel.

*Zeige mir die Fakten für alle Produkte im 4. Quartal 2003 und für Karlsruhe-Stadt und -Land.*

