

1. Übungsblatt

Ausgabe: 22. April 2020 · Besprechung: ab. 29. April 2020

Allgemeine Hinweise

- Teile der Übung sind an das Buch

J. Mundy, W. Thornthwaite, R. Kimball: The Microsoft Data Warehouse Toolkit (2nd edition), Wiley 2011,

angelehnt. Sie können es über die Unibibliothek als E-book beziehen.

Aufgabe 1: Online Transactional Processing (OLTP)

Wir betrachten einen Geschäftsprozess, der in einer Firma eingesetzt wird um Bestellungen zu verarbeiten. Der Geschäftsprozess beinhaltet einige Teilschritte, die in Abbildung 1 dargestellt werden. Der grobe Ablauf gestaltet sich folgendermaßen:

Erhält die Firma Kaufanfragen von Kunden werden dazu entsprechende Angebote generiert und an die Kunden verschickt. Nimmt der Kunde das Angebot an, muss er dies durch eine Vorauszahlung bestätigen. Dann erzeugt das IT-System einen Auftrag zur Produktion der Waren. Ist der Auftrag erfüllt werden die Waren geliefert und der Verkauf wird abgerechnet. Die Bestellung ist abgeschlossen.

Die Geschäftsprozesse der Firma werden elektronisch gesteuert, somit werden die getätigten *Transaktionen*, wie z.B. das Versenden von Angeboten oder Zahlungseingänge, in einer Datenbank erfasst.

a) Entity-Relationship Diagramm

Zeichnen Sie ein *Entity-Relationship Diagramm (ER Diagramm)* für die Daten, die dem Geschäftsprozess aus Abbildung 1 zugrunde liegen. Gehen Sie dabei folgendermaßen vor:

Bestimmen Sie zunächst die Entitäten, die am Geschäftsprozess beteiligt sind (z.B. Order, Product, Customer) und zeichnen Sie Boxen für die Entitäten. Als nächstes überlegen Sie sich in welcher Beziehung die Entitäten zueinander stehen. Tragen Sie die Beziehungen als Kanten zwischen den Entitäten ein (z.B. Angebot enthält 1 bis * Produkte).

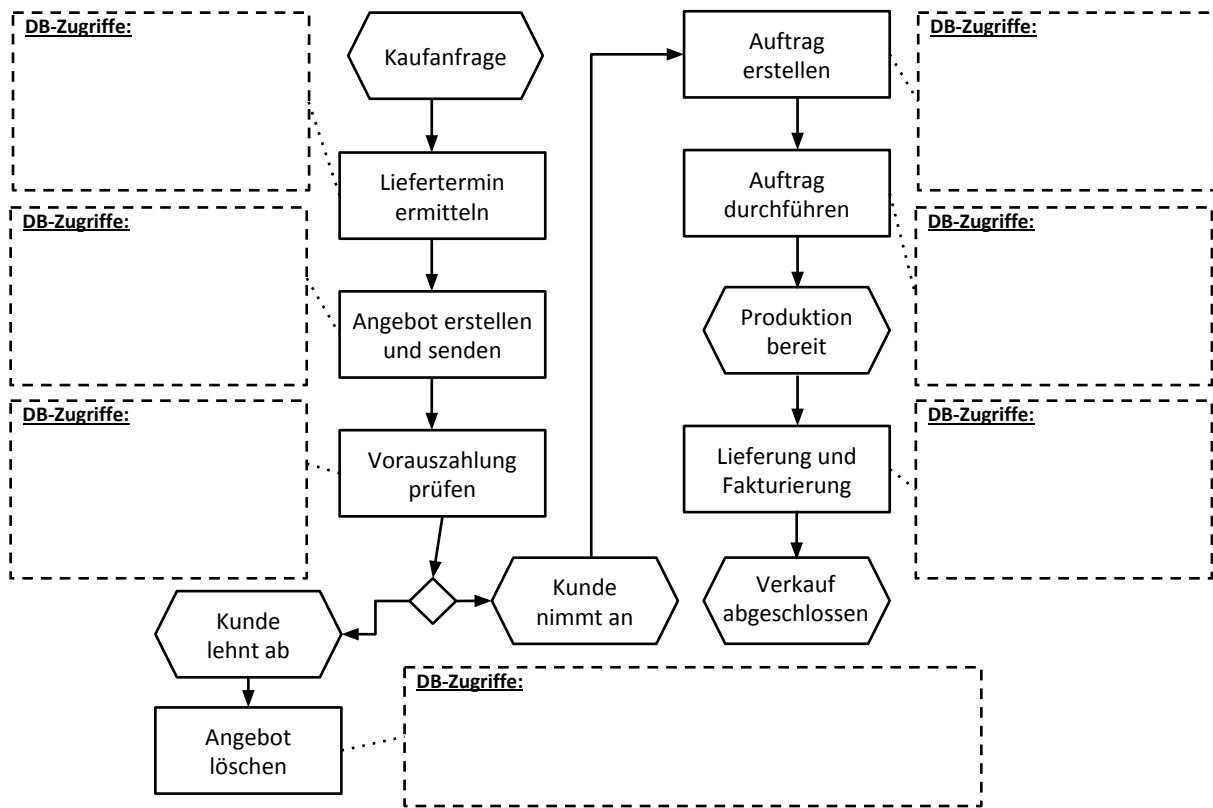


Abbildung 1: Beispielhafter, vereinfachter Ablauf eines Verkaufsprozesses

Table: OrderItem

OrderItemKey	ProductKey	OrderKey	Quantity
5	1500	239	30
6	2311	239	1
7	1800	239	7
8	1500	240	20
9	900	240	150
10	1800	241	4
...			

Abbildung 2: DB-Tabelle zur Speicherung von Daten des Geschäftsprozesses.

b) Datenbankzugriffe

Im praktischen Einsatz werden die Elemente des ER Diagramms typischerweise in *Tabellen* überführt. Als Beispiel zeigt Abbildung 2 den Auszug einer Tabelle `OrderItem`, die Einträge für alle Bestellposten enthält. Wir bezeichnen die Zeilen der Tabelle als *Tupel*. Die Tupel werden über `OrderItemKey` identifiziert; `ProductKey` ordnet ihnen ein Produkt zu und `OrderKey` eine Bestellung. `Quantity` gibt die Stückzahl in der Bestellung an.

Aufgabe:

Schätzen Sie wie viele Tupel bei der Durchführung des Geschäftsprozesses *gelesen* und *geschrieben* werden müssen. Tragen Sie die Anzahl der Zugriffe für jeden Teilschritt des Geschäftsprozess in das Feld DB-Zugriffe ein. Geben Sie jeweils die entsprechenden Tabellen an.

Beispiel:

Liefertermin ermitteln für eine Bestellung mit 5 Artikeln. Zunächst werden die Bestellposten aus der Tabelle `OrderItem` gelesen (5 Lesezugriffe). Danach wird je Bestellposten der Liefertermin abgefragt (5 weitere Lesezugriffe). Zuletzt wird der Liefertermin für die gesamte Bestellung in die Tabelle für Angebote geschrieben (1 Schreibzugriff).

b) Datenbankzugriffe

Überlegen Sie sich welche SQL Operationen während der Transaktionen ausgeführt werden.

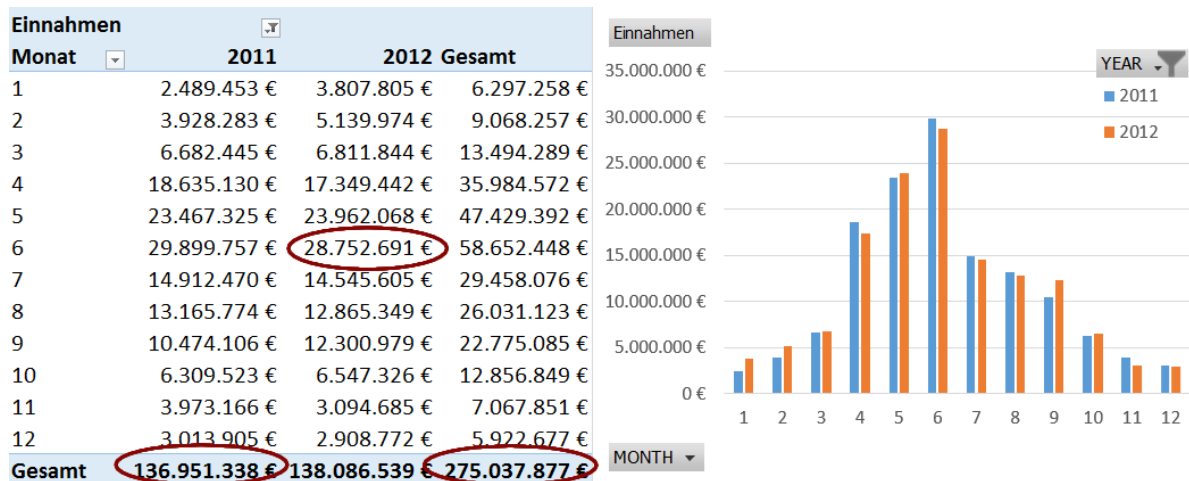


Abbildung 3: Analyse Report

Aufgabe 2: Online Analytical Processing (OLAP)

Der Datenbestand aus vergangenen Transaktionen wird häufig zur Durchführung von Geschäftsanalysen genutzt. Als Beispiel zeigt Abbildung 3 einen Analysebericht (*Report*) zu Verkäufen aus den Geschäftsjahren 2011 und 2012.

a) Analytische Anfragen

Welche Datenbankanfragen müssen ausgeführt werden um den Report aus Abbildung 3 mit Inhalt zu füllen? Überlegen Sie sich beispielhaft den SQL Code. Sie dürfen davon ausgehen, dass das Datenbankschema nur zwei Tabellen enthält. Die Tabelle *Date* enthält alle relevanten Zeitpunkte und *SalesItem* enthält die Bestellposten der getätigten Verkäufe.

b) Datenbankzugriffe

Im Bericht sind drei Werte mit einem roten Oval markiert. Bei der Erstellung des Reports wurde jeder der Werte als Ergebnis einer Datenbankanfrage ermittelt. Geben sie jeweils eine *Schätzung* an für die Anzahl der Tupel die zur Berechnung des Werts gelesen und geschrieben werden. Sie dürfen vereinfachend annehmen, dass je Monat 10.000 Produkte verkauft werden.

c) Vergleich von OLAP und OLTP

Überlegen Sie welche *Anforderungen* sich für den *Entwurf und Betrieb der Datenbanken* in den beiden Szenarien ergeben. Machen Sie sich die Anforderungen am Beispiel des Geschäftsprozesses aus Aufgabe 1 und des Reports aus Aufgabe 2 klar. Wie *unterscheiden* sich die Anforderungen für OLTP und OLAP? Diskutieren Sie, inwieweit die Anforderungen für OLAP und OLTP im *Konflikt* stehen (z.B. im Bezug auf Mehrbenutzerbetrieb und Konsistenzgarantien).