

Übungsblatt 4

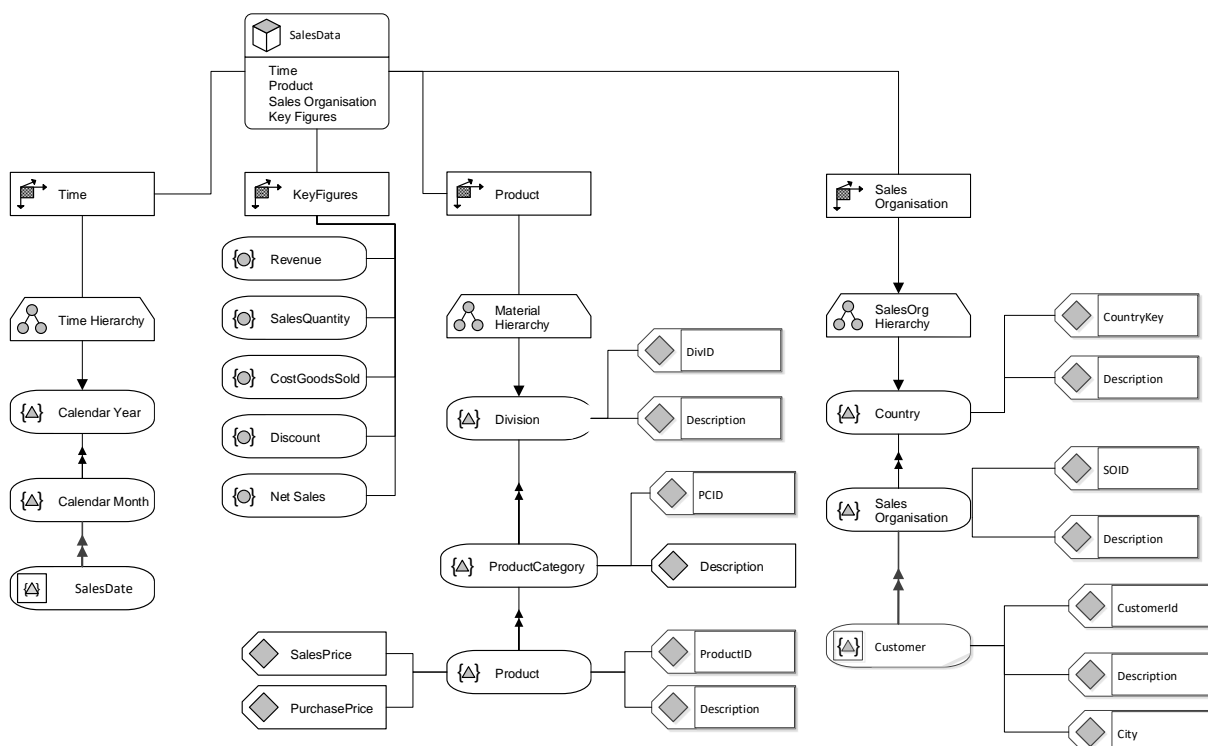
Stand: 08.11.2018

Bearbeiten Sie alle Aufgaben aller Übungsblätter – soweit nicht anders angegeben – in festen **Zweiergruppen**. Beide Gruppenmitglieder müssen die Aufgabe gemeinsam bearbeiten und die Lösung alleine erläutern können.

Aufgabe 6 (BI – Relationales Modell)

Nach der erfolgreichen Implementierung eines ERP-Systems im gesamten Dirt Bikes Unternehmen, wird im Management der Wunsch geäußert, ein einheitliches Reporting aufzusetzen für das Geschäft in den USA und in Europa.

Als Pilotbereich wurde der Verkauf (Sales) ausgewählt. Zusammen mit einem Beratungsunternehmen wurden durch Befragung des Managements die relevanten Anforderungen identifiziert. Die Berater haben im nachfolgenden ADAPT-Modell die Kernergebnisse zusammengefasst.

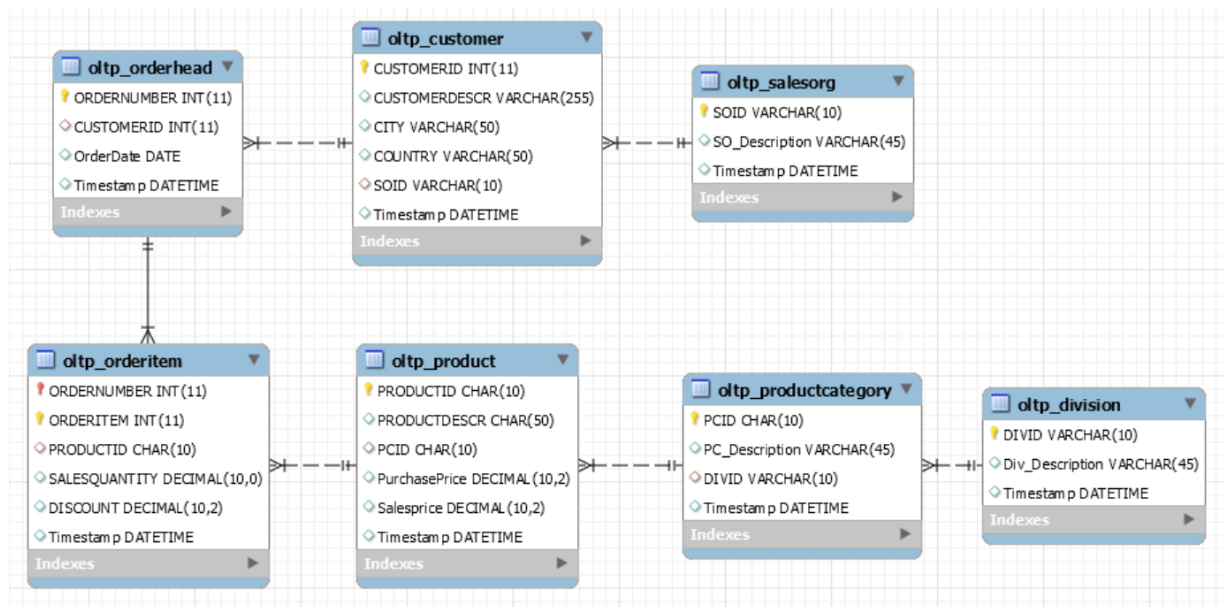


Sie sind jetzt mit der technischen Umsetzung beauftragt. Übertragen Sie das obige Modell in ein **Star-oder Snowflakeschema**. Legen Ihre Tabellen in der Oracle DB der HAW in Ihrem Tablespace an.

Abzugeben sind das Datenmodell Ihres Star- bzw. Snowflakeschemas.

Aufgabe 7 (BI - ETL)

Ihre Tabellen sollen nun mit den Daten des OLTP-Systems (ERP) befüllt werden. Das ERP-System basiert auf einer Oracle-Datenbank (Oracle Datenbank der HAW). Im **Schema ach053** sind die entsprechenden Tabellen gespeichert (*sie können also z.B. über `select * from ach053.oltp_orderhead` auf die Daten zugreifen*). Folgendes Datenmodell wird im ERP-System zur Speicherung der Kundenaufträge (transaktionalen Daten) verwendet.



Erstellen Sie mit einer ETL-Software (z.B. Pentaho Data Integration (**Community Edition**)) einen ETL-Prozess, der die Dimensionstabellen und die Faktentabelle Ihres in Aufgabe 6 angelegten Schemas initial befüllt.

Während des Befüllens müssen die gewünschten Kennzahlen berechnet werden. Auf folgende Berechnungsvorschriften hat sich das Management geeinigt.

- Revenue = SalesQuantity * Product.SalesPrice
- CostGoodsSold = SalesQuantity * Product.PurchasePrice
- Discount = Discount (siehe oltp_orderitem)
- NetSales = Revenue – Discount

Inkrementelles Laden: Implementieren Sie Ihren Prozess so, dass bei der erneuten Ausführung (inkrementelles Laden) nur neue Datensätze in die Faktentabelle eingefügt werden (*Bei den Tupeln der Dimensionstabellen sind keine Veränderungen zu erwarten*).

Überprüfen Sie Ihre Befüllung anhand der folgenden Zahlen.

		Orderdate				
Country		2013	2014	2015	2016	2017
DE	Netsales	142.471.116	144.200.795	141.806.086	147.607.853	147.023.236
	Salesquantity	36.308	36.554	35.948	37.538	37.530
US	Netsales	150.819.363	140.126.272	103.406.605	105.193.839	103.459.334
	Salesquantity	38.766	35.622	26.276	26.936	26.230

Abzugeben Datei, die Definition Ihres ETL-Prozesses enthält (z.B. bei pentaho die ktr-Datei)

Oracle Verbindungsdaten:

Host Name:

ora14.informatik.haw-hamburg.de

Database Name:

inf14

Tablespace for Data

ihre a-Kennung

Tablespace for Indices

Port Number:

1521

Username:

Ihre a-Kennung

Password:

••••••••

JDBC-Treiber: Der JDBC-Treiber für Oracle liegt in EMIL bei der Aufgabenstellung

Hinweise Pentaho (falls Sie pentaho verwenden wollen)

- Pentaho läuft nur bis Java 8 (Java 10 wird noch nicht unterstützt)
- Wenn Sie **Änderungen an den Tabellendefinitionen** vornehmen während der Erstellung des ETL-Prozesses müssen Sie pentaho zur Aktualisierung der Metadaten zwingen. Dies erfolgt im Menü unter **Tools → Database → Clear Cache**

Allgemeine Hinweise (gelten für alle Übungsblätter)

Ich akzeptiere nur Lösungen, auf denen **die Namen beider Bearbeiter, die Veranstaltungsbezeichnung, die Nummer des Übungsblattes und die Nummer der Praktikumsgruppe (1/2/3/4) notiert ist**. Außerdem müssen sie eine **Versionsnummer** auf dem Lösungsblatt notieren und die Versionen bitte bei erneuter Abgabe hochzählen.

Je Zweier-Gruppe ist in EMIL nur von einem Gruppenmitglied das Lösungsdokument hochzuladen. Die Abgabe der Lösungen muss spätestens am **Tag der Übung in EMIL** erfolgen (**Deadline: Ende des jeweiligen Praktikumstermins**).

Sind Lösungen von zwei/ mehreren Gruppen zu ähnlich, bekommen alle beteiligte Gruppen einen Fehlversuch und alle Gruppen müssen eine überarbeitete Version einreichen.