Laboratory 06
Level of Aggregation

Case Study 1 – Clothing Company

Let's consider a Clothing case study. The operational database consists of the following tables:

CUSTOMER1

CUSTID	LNAME	FNAME	ADDRESS	PHONE	CITY
107	Smith	John	731 Plenty Road	9231455	Clayton
232	Wong	Franklin	638 Voss Street	9756945	Preston
133	Zelaya	Alicia	3321 Castle Ave	9867055	Balwyn
154	Wallace	Jennifer	291 Berry Street	9234536	Preston
179	Narayan	Ramesh	975 Fire Road	9456738	Carlton
181	Jane	Adam	229 Clayton Road	9543877	Clayton
183	Judy	Backhouse	122 Rose Street	9235345	Caulfield

ORDER1

ORDERID	ORDERDATE	PAYMETHOD	ORDERSOURCE	CUSTID
1057	20/02/06	CARD	WEB SITE	107
1058	03/03/06	CARD	PHONE	232
1059	12/03/06	CHEQUE	WEB SITE	133
1060	20/03/06	CHEQUE	WEB SITE	133
1061	10/04/06	CARD	FAX	179
1062	01/04/06	CARD	FAX	179
1063	07/09/06	CARD	WEB SITE	154
1064	14/07/06	CARD	WEB SITE	154
1065	30/11/06	CARD	PHONE	179
1066	20/01/06	CHEQUE	WEB SITE	179

ORDER_INV1

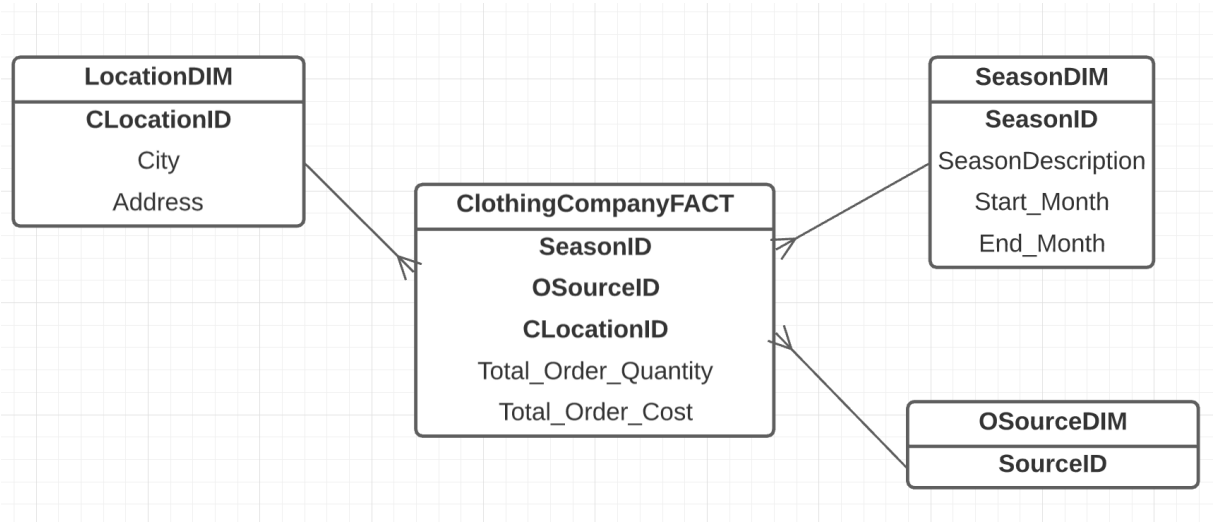
ORDERID	INVID	ORDERPRICE	QUANTITY
1057	11668	259.99	10
1058	11668	239.99	20
1059	11780	21.99	5
1060	11776	20.99	50
1061	11779	29.95	25
1061	11780	29.95	50
1062	11669	229.99	40
1063	11778	25.95	50
1064	11779	29.95	12
1065	11780	26.95	32
1066	11775	29.95	30

INVENTORY1 (*QOH = Quantity on Hand*)

INVID	QOH	ITEMID	ITEMSIZE	COLOUR
11668	16	786	M	Sienna
11669	12	786	L	Forest
11775	150	894	S	Khaki
11776	147	894	M	Khaki
11777	0	894	L	Khaki
11778	139	894	S	Olive
11779	137	894	M	Olive
11780	115	894	L	Olive

ITEM1			
ITEMID	**CURRENT PRICE**	**ITEMDESC**	**CATEGORY**
894	29.95	Women's Hiking Shorts	Women's Clothing
897	200.95	Women's Fleece Pullovers	Women's Clothing
995	50.00	Children's Beachcomber Sandals	Children's Clothing
559	35.00	Men's Expedition Parka	Men's Clothing
786	259.99	3-Season Jacket	Men's Clothing

A star schema has been created:



The fact measures included in the above star schema are TotalOrderQuantity, which is taken from the Quantity attribute from table **ORDER_INV1**, and TotalOrderCost, which is based on order price * quantity.

The above star schema contains highly aggregated data, and let's assume that this star schema is at level-2 in the data warehouse architecture.

Questions: Draw and implement level-1 and level-0 star schemas for the above clothing data warehouse case study.

The following operational databases have been provided for you:
clothing.Customer1: table that stores information about customer.
clothing.Inventory1: table that stores information about inventory.
clothing.Item1: table that stores information about item.
clothing.Order_Inv1: table that stores information about order and inventory.
clothing.Order1: table that stores information about order.

Case Study 2 – Toll Way

There is a toll way in a metropolitan city (such as CityLink or EastLink in Melbourne, or any similar toll roads in other major cities in the world). This toll way has a number of gates, where the motorist needs to pay. Every time a motorist passes through this toll gate, the registration number of the vehicle, vehicle type (e.g. car, bus, truck, etc), amount paid, and time, are recorded in the operational database.

A data warehouse needs to be built, for analysing the *revenue* from the toll payments. The management would like to drill down this revenue based on the *tollgate* (there is a number of toll gates along the toll way), *day of week* (e.g. weekdays, weekends), and *time period of a day* (e.g. peak hours, non-peak hours, late nights).

You are required to draw **three levels of star schemas** showing three different levels of aggregation for the above data warehouse. Level-0 star schema contains the most detailed data, whereas level-2 star schema is the highly aggregated (e.g. containing highly aggregated data).

Questions:

(a) Draw a level-2 star schema.

(b) Draw a level-1 star schema. You may want to add a new dimension, called *vehicle* (e.g. cars, trucks, busses, etc).

(c) Draw a level-0 star schema.

THE END