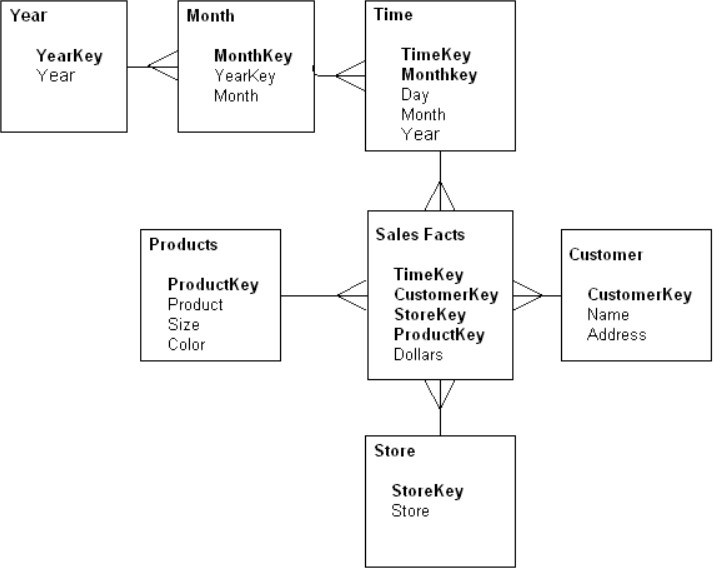
**DATAWARE HOUSE ASSESMENT-1**

**1:For the given Dimensional Modelling , please identify the following**

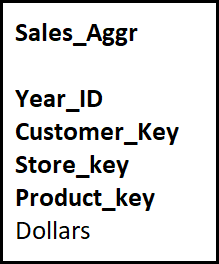
* **How many dimensions and facts are present?**

In the above given dimensional modelling we have one(1) fact table that is "sales fact" and we have six(6) dimensions table that is (Year,Month,Time,Products,customer,Store) are present in the above dimensional model.

* **Please identify the cardinality between each table?**

|  |  |
| --- | --- |
| Dimension Tables | Cardinality Between each Table |
| Year-to-Month | Cardinality between Year-to-Month is One-to-Many |
| Month-to-Time | Cardinality between Month-to-Time is One-to-Many |
| Time-to-Sales Facts | Cardinality between Time-to-Sales Facts is One-to-Many |
| Products-to-Sales Facts | Cardinality between Products-to-Sales Facts is One-to-Many |
| Customer-to-Sale Facts | Cardinality between Customer-to-Sales Facts is One-to-Many |
| Store-to-Sale Facts | Cardinality between Store-to-Sales Facts is One-to-Many |

* **How to create a Sales\_Aggr fact using the following structure(SQL Statement):**



CREATE TABLE Sales\_Aggr as (select year\_key as “year\_id”,customerKey as “customer\_key”,StoreKey as “Store\_Key”,ProductKey as “Product\_Key” from year,Products,Customer,Store);

ALTER TABLE Sales\_Aggr ADD Dollars double(40);

* **Can you please Modify the above snowflake schema to star schema and draw the dimension model,showing all the cardinality?**

**Date**

**Datekey**

Day

Month

Year

**Store**

**StoreKey**

Store

**Sales Facts**

**Datekey**

**CustomerKey**

**StoreKey**

**ProductKey**

Dollars

**Customer**

**CustomerKey**

Name

Address

**Products**

**ProductKey**

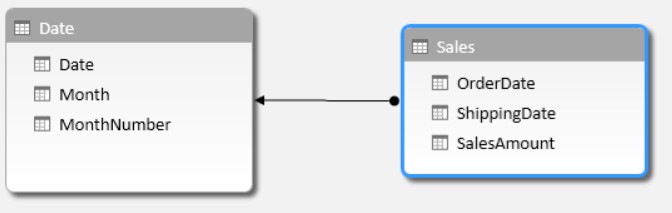
Product

Size

Color

|  |  |
| --- | --- |
| Dimension Tables | Cardinality Between each Table |
| Date-to-Sales Facts | Cardinality between Date-to-Sales Facts is One-to-Many |
| Products-to-Sales Facts | Cardinality between Products-to-Sales Facts is One-to-Many |
| Customer-to-Sale Facts | Cardinality between Customer-to-Sales Facts is One-to-Many |
| Store-to-Sale Facts | Cardinality between Store-to-Sales Facts is One-to-Many |

**2:For the following dimension Model can you please give an example of Circular Join and how to avoid it:**



**DATE**

|  |  |  |
| --- | --- | --- |
| Date | Month | MonthNumber |
| 05-02-2019 | FEB | 02 |
| 14-08-2019 | AUG | 08 |
| 23-04-2019 | APR | 04 |

**SALES**

|  |  |  |
| --- | --- | --- |
| OrderDate | ShippingDate | SalesAmount |
| 15-02-2019 | 20-02-2019 | 2000 |
| 24-08-2019 | 30-08-2019 | 4000 |
| 10-04-2019 | 18-04-2019 | 6000 |

**In the above table following Query will Create the circular Join**

SELECT SAL.OrderDate,SAL.ShippingDate

FROM DATE DA,SALES SAL

WHERE

DA.DATE=SAL.OrderDate AND

DA.DATE =SAL.ShippingDate;

**In the above table we can remove circular join by using alias to the attributes**

SELECT SAL.SalesAmount,SAL.OrderDate,SAL.ShippingDate

FROM Date AS “OrderDate”,Date as “ShippingDate”,

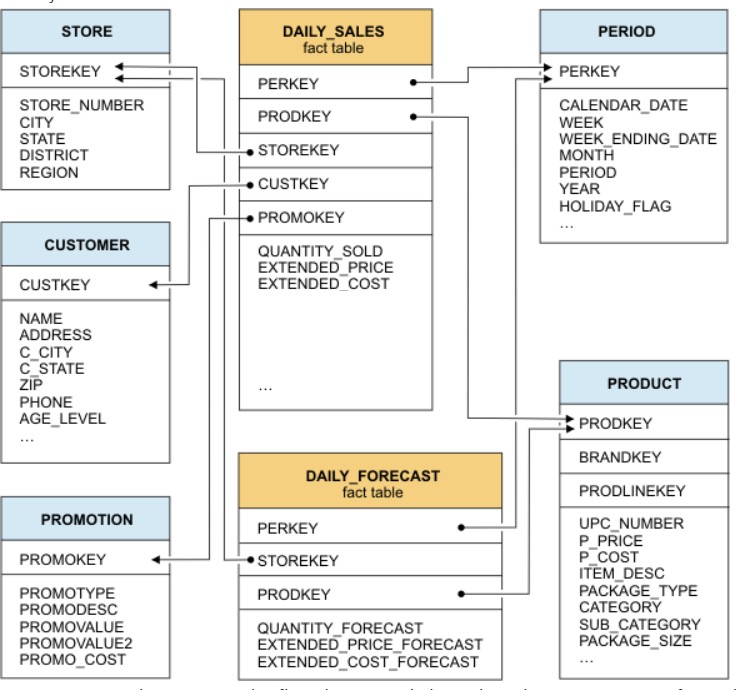
SALES SAL,DATE DAT

WHERE

OrderDate.DATE=SAL.OrderDate AND

ShippingDate.DATE=SAL.ShippingDate;

**3:For the given Dimension Model, can you please generate a sql to get the total divergence between Quantity sold and Quantity Forecast for the current month for all the stores:**

****

SELECT Sum(QUANTITY\_SOLD) - Sum(QUANTITY\_FORECAST) as “DIVERGENCE” FROM

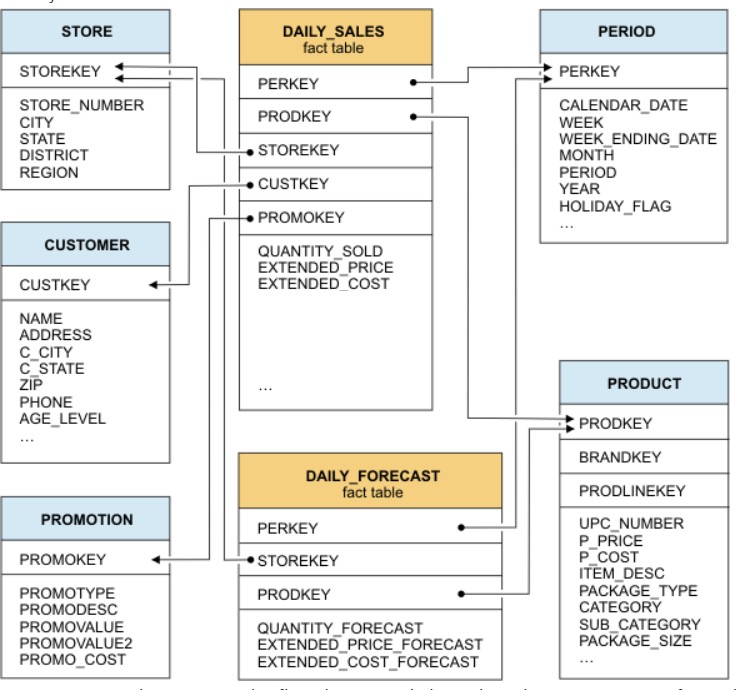
DAILY\_SALES,DAILY\_FORECAST,PERIOD PER

WHERE

Month(PER.Month)-Month(Current\_Date)

GROUP BY STORE\_KEY;

**4:For the above-mentioned dimension model, please identify the conformed and non- conformed dimensions. Additionally, identify the measure types?**

****In the above diagram shows that

**CONFORMED DIMENSIONS:**

* STORE
* PERIOD
* PRODUCT

**NON-CONFORMED DIMENSIONS:**

* PROMOTION
* CUSTOMER

**MEASURES**

* ADDITIVE:
* QUANTITY\_SOLD
* QUANTITY\_FORECAST
* SEMI-ADDITIVE:
* EXTENDED\_PRICE
* EXTENDED\_COST
* EXTENDED\_PRICE\_FORECAST
* EXTENDED\_COST\_FORECAST
* NON-ADDITIVE:
* In this model there is no Non-Additive Measures because it does not consist of any percentages or ratios are not calculated

**5:Make a list of differences between DW and OLTP based on Size, Usage, Processing and Data Models.**

|  |  |  |
| --- | --- | --- |
|  | **DATA WAREHOUSE** | **ONLINE TRANSCATION PROCESSING** |
| **SIZE** | Size of Data Warehouse  is 10MB-100GB | Size of Online Transaction Processing is 100GB-2TB |
| **USAGE** | Data Warehouse uses repetitive usage | Online Transaction Processing uses ad-hoc usage |
| **PROCESSING** | Data Warehouse uses  Query Processing | Online Transaction uses  Transaction Processing |
| **DATA MODELS** | Data Warehouse uses E-R Modeling | Online Transaction Processing uses Dimensional Modeling |