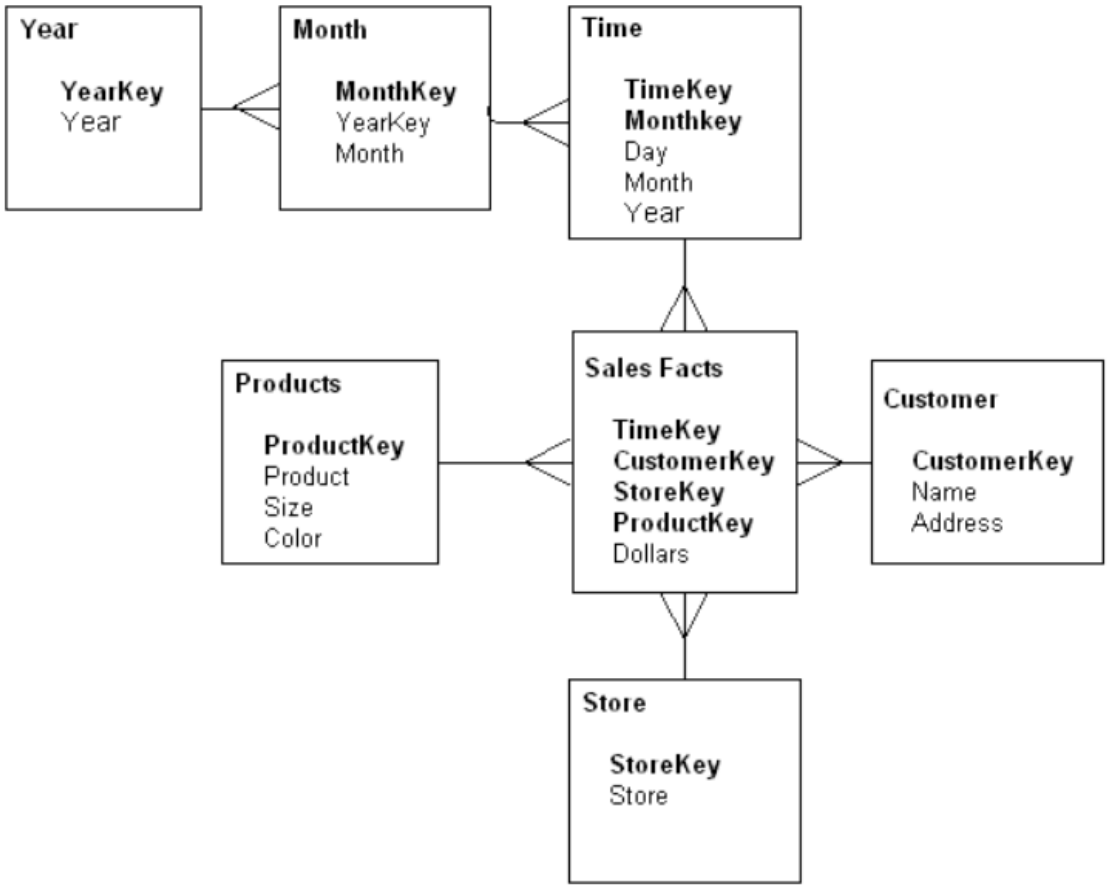
Syed Mushtaq DATA WAREHOUSE ASSESMENT

Q1



Q1.1 . How many Facts and Dimensions are present?

Ans . FACTS =1 **DIMENSIONS=6**

Facts	**Dimensions**
SALES	CUSTOMER
	STORE
	PRODUCTS
	TIME
	MONTH

YEAR

Q1.2 Please identify the cardinality between each table ?

Ans Cardinality refers to the maximum number of times an instance in one entity can relate to instances of another entity.

Tables	**Relations**
year - month	1 : M
Month - Time	1 : M
Time - Sales Facts	1 : M
Products - Sales Facts	1 : M
Customer - Sales Facts	1 : M
Store - Sales Facts	1 : M

Q1.3 How to create Sales_Aggr Fact useing the following Structure (SQL Statement) ?

CREATE TABLE Sales_Aggr (

Year_ID int(15),Customer_Key int(15),

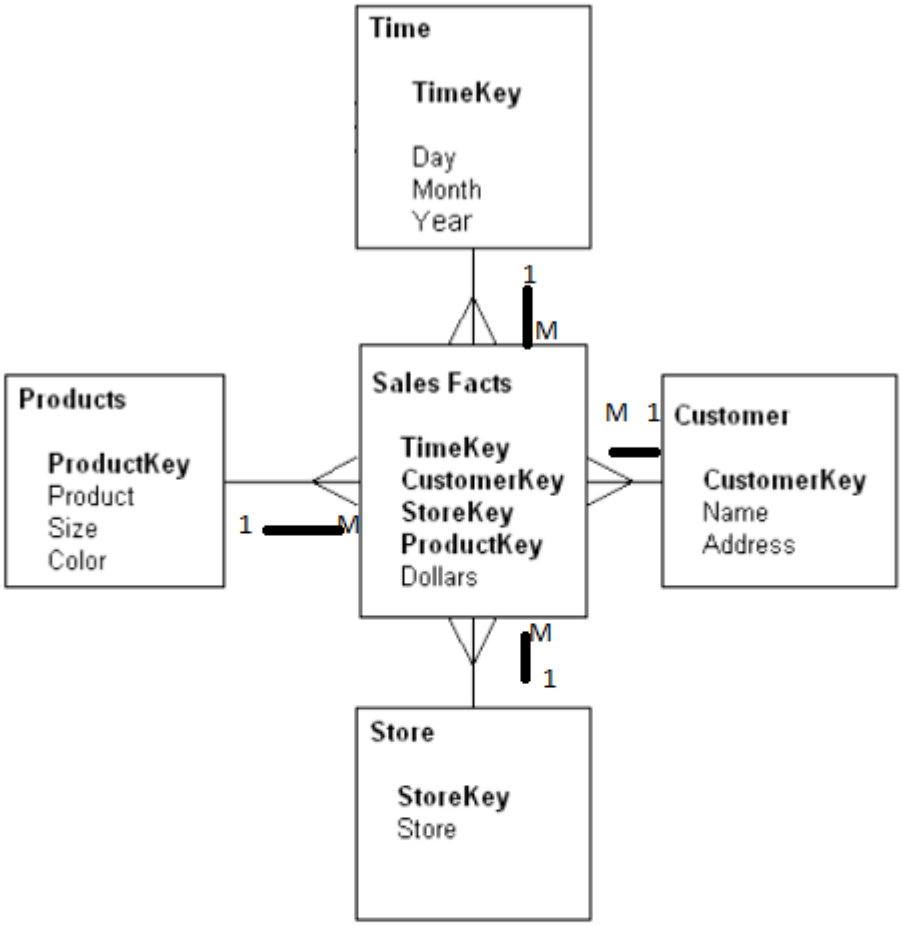
Store_Key int(15),Product_Key int(15),

```
Dollars Double     ,

Foreign Key (Year_ID ) References Year(YearKey),

Foreign Key (CustomerKey) References
Year(CustomerKey),

Foreign Key (Store_Key ) References Year(StoreKey),

Foreign Key (Product_Key ) References
Year(ProductKey) , Primary
Key(Year_ID,Store_Key,Customer_Key,Product_Key));

----------------------------------------------------------------

INSERT INTO Sales_Aggr Values ( select Y.YearKey,

  C.CustomerKey ,S.StoreKey,P.ProductKey
,sum(P.Price*S.ProductKey) FROM Year Y ,Customer
C,Store S,Products P

GROUP BY Y.YearKey,

C.CustomerKey ,S.StoreKey,P.ProductKey;
```

Q1.4 Can you please Modify the above Snowflake schema to Star schema and draw the dimensional model , showing all the cardinality ?

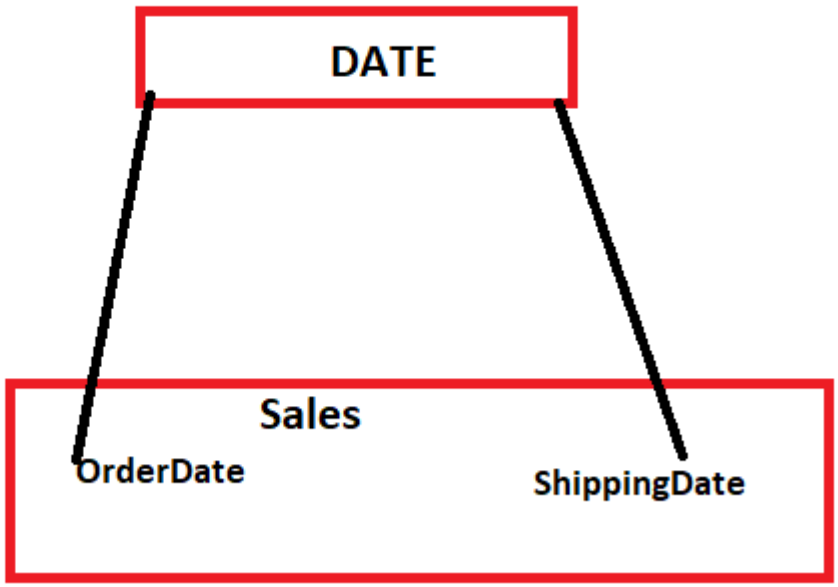
Ans SELECT TimeKey, CustomerKey, StoreKey , ProductKey , Dollars FROM

Sales_Facts ;



Q2 . For the following dimension Model can you please give an example of Circular join and how to avoid it:

Ans SELECT SUM(SalesAmount)
from Sales s ,Date d
where s.OrderDate = d.Date
AND
s.ShippingDate = d.Date;

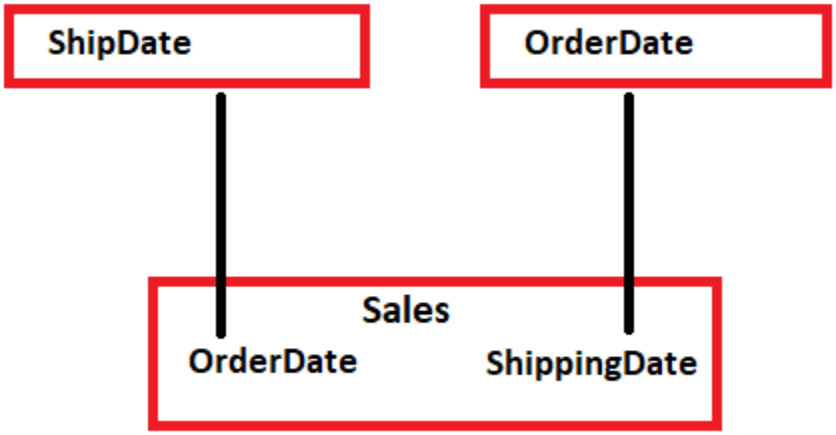


The above query give circular join cause both **OrderDate** and **ShippingDate** will point to " same" **Date.**

--

Avoiding Circular Join

```
SELECT    SUM(SalesAmount)
from Sales s ,Date orderdate, Date shipdate
where s.OrderDate = orderdate.Date
AND
s.ShippingDate = shipdate.Date;
```

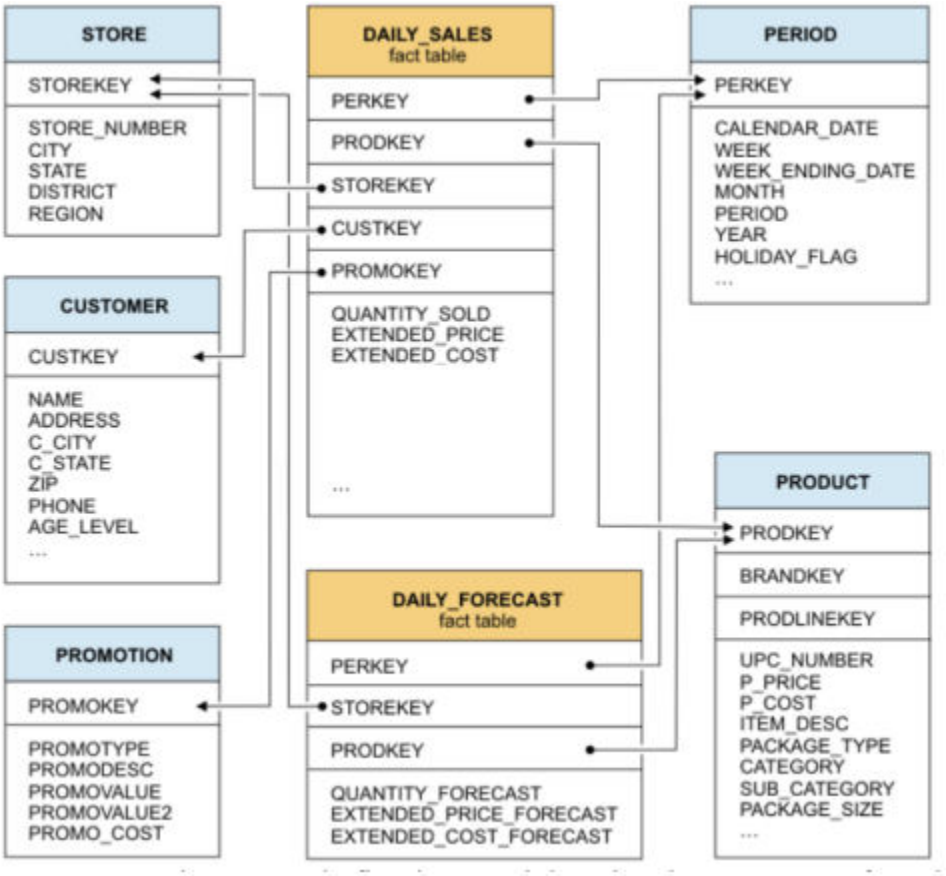


Q 3 . For the given Dimension Model, can you please generate a sql to get the total
divergence between Quantity sold and Quantity Forecast for the current month for
all the stores:

Ans .

```
        SELECT SUM(DS.QUNATITY_SOLD) -SUM(DF.QUANTITY_FORECAST)
as DIVERGENCE    FROM      PERIOD    P
  INNER JOIN DAILY_SALES    DS ON DS.PERKEY = P.PERKEY
    INNER JOIN DAILY_FORECAST    DF ON DF.PERKEY = P.PERKEY
```

AND WHERE P.MONTH = EXTRACT (MONTH FROM CURDATE()
AND WHERE P.YEAR =EXTRACT (YEAR FROM CURDATE());



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

Ans

Conformed	**Non-Conformed**

Store	Customer
Product	Promotion
Period	

Quantity_Sold	**Additive Measure**
Quantity_Forecast	**Additive Measure**
Extended_Price	**Semi - Additive Measure**
Extended_Price_Forecast	**Semi - Additive Measure**
Extended_Cost	**Semi - Additive Measure**
Extended_Cost_Forecast	**Semi - Additive Measure**

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

Ans

Specification	**Data Warehouse**	**OLTP**
Processing	Depends on the amount of data involved ;batch data refreshes and complex queries may take many hours. Cause it is Normalised .	Very Fast. Cause it is De -Normalised .
Usage	To help with planning ,problem solving,and decision support	To control and run fundamental business

		tasks
Size	Can be relatively small if historical data is archived	Large due to existence of Aggregation structures and history data
Data Models	Dimensional Modelling	ER Modelling