DATAWARE HOUSING AND DATA MINING

ASSIGNMENT 3

GROUP MEMBERS

MUHAMMAD BILAL AHMED (FA20-BCS-041)

MUHAMMAD SALMAN (SP20-BCS-131)

1. **Datawarehouse creation**

We created Datawarehouse in SQL server using SQL Queries.

Datawarehouse consist of

1. **Dimensions**

Dimensions are individual, non-overlapping elements. The dimensions facilitate the filtering, grouping, and labelling of the data. Dimension tables have textual descriptions about the business subjects. In our case Dimensions were Customer, Employees, Time, Products.

1. **Fact Table**

The data stored in a Fact Table is known as Measures (also dependent attributes). The fact table should give statistics of sales broken down by customer, product, period, Employees. Fact table contains following attributes

* **Foreign keys:**

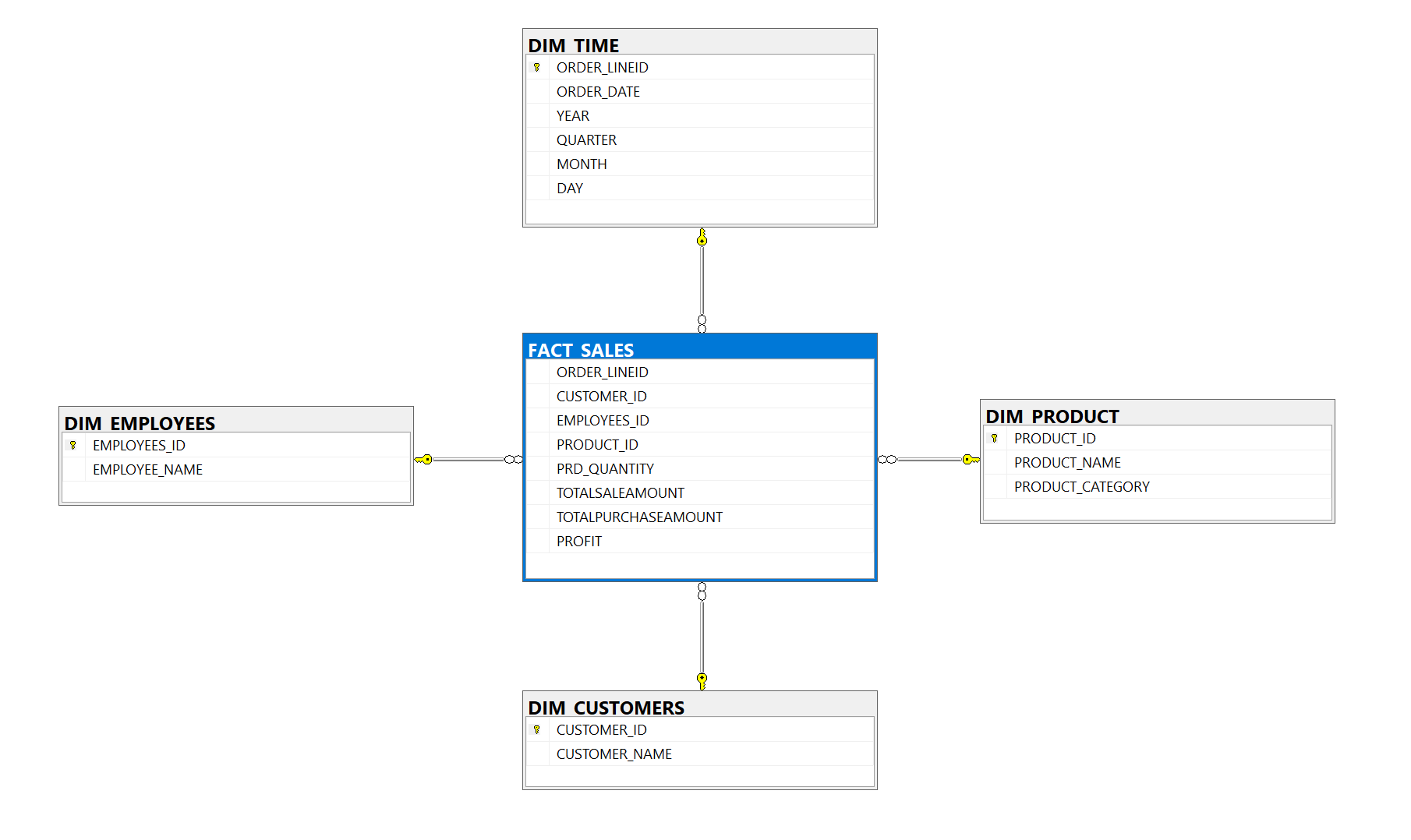
Customer id, Employee id, Order line id, Product id

* **Measure:**

Total purchase amount, Total sale amount, profit

We use star schema for dimension modeling.

**Star schema:**

****

1. **Creating Dimensions and inserting data using SQL SERVER**

**Product Dimension**

CREATE TABLE DIM\_PRODUCT

(

PRODUCT\_ID INT NOT NULL PRIMARY KEY,

PRODUCT\_NAME VARCHAR(50) NOT NULL,

PRODUCT\_CATEGORY VARCHAR(50) NOT NULL,

SALE\_PRICE INT NOT NULL,

);

**Inserting data into Product Dimension**

INSERT INTO DIM\_PRODUCT(PRODUCT\_ID, PRODUCT\_NAME, PRODUCT\_CATEGORY, SALE\_PRICE)

SELECT DISTINCT (PRODUCTS.PRODUCT\_ID),

PRODUCTS.PRODUCT\_NAME, PRODUCT\_CATEGORY.CATEGORY\_NAME, [DWDM\_PROJECT].[DBO].ORDER\_LINE.PRICE\_EACH

FROM [DWDM\_PROJECT].[DBO].PRODUCTS

JOIN [DWDM\_PROJECT].[DBO].PRODUCT\_CATEGORY

ON PRODUCTS.PRD\_CATEGORY\_ID = PRODUCT\_CATEGORY.PRD\_CATEGORY\_ID

JOIN [DWDM\_PROJECT].[DBO].ORDER\_LINE

ON PRODUCTS.PRODUCT\_ID = [DWDM\_PROJECT].[DBO].ORDER\_LINE.PRODUCT\_ID

ORDER BY PRODUCTS.PRODUCT\_ID ASC;

**Customer dimension**

CREATE TABLE DIM\_CUSTOMERS

(

CUSTOMER\_ID INT NOT NULL PRIMARY KEY,

CUSTOMER\_NAME VARCHAR(50),

);

**Inserting data in Customer dimension**

INSERT INTO DIM\_CUSTOMERS(CUSTOMER\_ID, CUSTOMER\_NAME)

SELECT CUSTOMERS\_ID,CUSTOMER\_NAME FROM [DWDM\_PROJECT].[DBO].CUSTOMERS

**Time Dimension**

CREATE TABLE DIM\_TIME

(

ORDER\_LINEID INT NOT NULL PRIMARY KEY,

ORDER\_DATE DATE,

YEAR INT,

QUARTER INT,

MONTH INT,

DAY INT,

);

**Inserting data into Time Dimension**

INSERT INTO DIM\_TIME(ORDER\_LINEID, ORDER\_DATE, YEAR, QUARTER, MONTH, DAY)

SELECT [DWDM\_PROJECT].[DBO].ORDER\_LINE.ORDERLINE\_ID, [DWDM\_PROJECT].[DBO].ORDERS.ORDER\_DATE,

YEAR([DWDM\_PROJECT].[DBO].ORDERS.ORDER\_DATE) AS Year,

DATEPART(qq, [DWDM\_PROJECT].[DBO].ORDERS.ORDER\_DATE) AS QUARTER,

MONTH([DWDM\_PROJECT].[DBO].ORDERS.ORDER\_DATE) AS MONTH,

DATEPART(dd, [DWDM\_PROJECT].[DBO].ORDERS.ORDER\_DATE) AS DAY

FROM [DWDM\_PROJECT].[DBO].ORDER\_LINE

JOIN [DWDM\_PROJECT].[DBO].ORDERS ON

[DWDM\_PROJECT].[DBO].ORDER\_LINE.ORDER\_ID = [DWDM\_PROJECT].[DBO].ORDERS.ORDER\_ID

**Employee Dimension**

CREATE TABLE DIM\_EMPLOYEES

(

EMPLOYEES\_ID INT NOT NULL PRIMARY KEY,

EMPLOYEE\_NAME VARCHAR(100),

)

**Inserting data Employee Dimension**

INSERT INTO DIM\_EMPLOYEES(EMPLOYEES\_ID, EMPLOYEE\_NAME)

SELECT EMPLOYEES\_ID,EMPLOYEE\_NAME FROM [DWDM\_PROJECT].[DBO].EMPLOYEES

**Fact\_Sales**

CREATE TABLE FACT\_SALES

(

ORDER\_LINEID INT FOREIGN KEY REFERENCES DIM\_TIME(ORDER\_LINEID),

CUSTOMER\_ID INT FOREIGN KEY REFERENCES DIM\_CUSTOMERS(CUSTOMER\_ID),

EMPLOYEES\_ID INT FOREIGN KEY REFERENCES DIM\_EMPLOYEES(EMPLOYEES\_ID),

PRODUCT\_ID INT FOREIGN KEY REFERENCES DIM\_PRODUCT(PRODUCT\_ID),

PRD\_QUANTITY INT,

TOTALSALEAMOUNT INT NOT NULL,

TOTALPURCHASEAMOUNT INT NOT NULL,

PROFIT INT

);

**Inserting data into Fact table**

SELECT DISTINCT [DWDM\_PROJECT].[DBO].ORDER\_LINE.ORDERLINE\_ID,[DWDM\_PROJECT].[DBO].ORDERS.CUSTOMERS\_ID,[DWDM\_PROJECT].[DBO].ORDERS.EMPLOYEES\_ID,

[DWDM\_PROJECT].[DBO].ORDER\_LINE.PRODUCT\_ID,

[DWDM\_PROJECT].[DBO].ORDER\_LINE.PRD\_QUANTITY,

[DWDM\_PROJECT].[DBO].ORDER\_LINE.PRICE\_EACH \* [DWDM\_PROJECT].[DBO].ORDER\_LINE.PRD\_QUANTITY AS [TOTAL SALE AAMOUNT],

[DWDM\_PROJECT].[DBO].PURCHASE\_DETAILS.PURCHASE\_PRICE \* [DWDM\_PROJECT].[DBO].ORDER\_LINE.PRD\_QUANTITY AS [TOTAL PURCHASE AMOUNT],

(([DWDM\_PROJECT].[DBO].ORDER\_LINE.PRICE\_EACH - [DWDM\_PROJECT].[DBO].PURCHASE\_DETAILS.PURCHASE\_PRICE) \* [DWDM\_PROJECT].[DBO].ORDER\_LINE.PRD\_QUANTITY) AS [TOTAL PROFIT]

FROM [DWDM\_PROJECT].[DBO].ORDER\_LINE

INNER JOIN [DWDM\_PROJECT].[DBO].ORDERS

ON [DWDM\_PROJECT].[DBO].ORDER\_LINE.ORDER\_ID = [DWDM\_PROJECT].[DBO].ORDERS.ORDER\_ID

INNER JOIN [DWDM\_PROJECT].[DBO].PURCHASE\_DETAILS

ON [DWDM\_PROJECT].[DBO].ORDER\_LINE.PRODUCT\_ID = [DWDM\_PROJECT].[DBO].PURCHASE\_DETAILS.PRODUCT\_ID

ORDER BY ORDERLINE\_ID ASC;