Assignment 2

**Task 1: - Understanding the Data**

1. Describe the data in hand in your own words.

In superstore database there are five tables of information cust\_dimen, orders\_dimen, prod\_dimen, shipping\_dimen and market\_fact etc.

1. cust\_dimen: - This table contain all the information related the customers like id, customer name, region and costumers’ segment etc.
2. orders\_dimen: - This table contain all the information related the customers order details like order date, order id and order priority etc.
3. prod\_dimen: - This table contain all the information related the product like product id, product category and sub-category etc.
4. shipping\_dimen: - This table contain all the information related the shipping of product like shipping mode, shipping id and shipping date etc.
5. market\_fact: - This table contain all the information related the market fact of product like their sales, discount, order quantity, profit, shipping cost and product base margin etc.
6. Identify and list the Primary Keys and Foreign Keys for this dataset provided to you (In case you don’t find either primary or foreign key, then specially mention this in your answer)

|  |  |  |
| --- | --- | --- |
| TABLE NAME | PRIMARY KEY | FOREIGN KEY |
| cust\_dimen | Cust\_id | None |
| orders\_dimen | Ord\_id | Nope |
| prod\_dimen | Prod\_id | Order\_id |
| shipping\_dimen | Ship\_id | Order\_id |
| market\_fact | Nope | * 1. Cust\_id   2. Ord\_id   3. Prod\_id   4. Ship\_id |

We can use order\_id from order\_dimen table as primary key to uniquely identify the order date only if we split the order\_dimen table and order\_id as foreign key in shipping \_dimen table.

**Task 2: - Basic & Advanced Analysis (Answers)**

create database superstore;

select \* from superstore.cust\_dimen;

select \* from superstore.market\_fact;

select \* from superstore.orders\_dimen;

select \* from superstore.prod\_dimen;

select \* from superstore.shipping\_dimen;

use superstore;

#Que\_1\_=============================================================================================

select Customer\_Name as `Customer Name`,

Customer\_Segment as `Customer Segment`

from superstore.cust\_dimen;

#Que\_2\_=============================================================================================

select \* from superstore.cust\_dimen order by 1 desc;

#Que 3 ==============================================================================================

select Order\_ID , Order\_Date from superstore.orders\_dimen where Order\_Priority = 'HIGH';

#Que 4 =============================================================================================

select sum(Sales) as total\_sales,avg(Sales) as avg\_sales from superstore.market\_fact;

#Que 5 =============================================================================================

select max(Sales) as max\_sale,min(Sales) as min\_sale from superstore.market\_fact;

#Que 6 =============================================================================================

select Region,count(Customer\_Name) as 'Number\_of\_Customers' from superstore.cust\_dimen group by Region

order by count(Customer\_Name) desc;

#Que 7 ==============================================================================================

select Region,count(customer\_name) as 'Number\_of\_Customers' from superstore.cust\_dimen group by Region

order by count(Customer\_Name) desc limit 1 ;

#Que 8 ================================================================================================

select Customer\_Name, count(\*) as num\_tables from

superstore.market\_fact s, superstore.cust\_dimen c, superstore.prod\_dimen p

where s.Cust\_id = c.Cust\_id and s.Prod\_id = p.Prod\_id and

p.Product\_Sub\_Category = 'TABLES' and c.Region = 'ATLANTIC'

group by Customer\_Name;

#Que 9 ==============================================================================================

select Customer\_Name,count(\*) as `no of small business owners` from superstore.cust\_dimen

where Customer\_Segment = 'SMALL BUSINESS' and Province = 'ONTARIO'

group by Customer\_Name;

#Que 10 ==============================================================================================

select Prod\_id,count(\*) as `no\_of\_products sold` from superstore.market\_fact group by Prod\_id

order by count(`no\_of\_products sold`) desc;

#Que 11 ==============================================================================================

select Prod\_id,Product\_Sub\_Category from superstore.prod\_dimen

where Product\_Category='FURNITURE' or Product\_Category='TECHNOLOGY'

group by Prod\_id;

#Que 12 =================================================================================================

select Product\_Category,Profit from superstore.market\_fact s,superstore.prod\_dimen p

where s.Prod\_id = p.Prod\_id

group by Product\_Category order by Profit desc;

#Que 13 =====================================================================================================

select Product\_Category,Product\_Sub\_Category,Profit from superstore.market\_fact s,superstore.prod\_dimen p

where s.Prod\_id = p.Prod\_id;

#Que 14 =====================================================================================================

select Order\_Date,Order\_Quantity,Sales from superstore.market\_fact s, superstore.orders\_dimen c

where s.Ord\_id = c.Ord\_id;

#Que 15 ======================================================================================================

select Customer\_Name from superstore.cust\_dimen

where Customer\_Name like '\_R%' and Customer\_Name like '\_\_\_D%';

#Que 16 ===================================================================================================

select c.Cust\_id,s.Sales,c.Customer\_Name,c.Region from superstore.market\_fact s,superstore.cust\_dimen c

where s.Cust\_id = c.Cust\_id and Sales between 1000 and 5000;

#Que 17 ======================================================================================================

select min(Sales) as `3rd highest salary`

FROM (

select Sales from superstore.market\_fact order by Sales desc limit 3

) as a;

#Que 18 =====================================================================================================

select Region,count(Ship\_id) as no\_of\_shipment,sum(Profit) as profit\_in\_each\_region from

superstore.cust\_dimen c,superstore.market\_fact s,superstore.prod\_dimen p

where c.Cust\_id = s.Cust\_id and s.Prod\_id = p.Prod\_id

group by Region

order by profit\_in\_each\_region asc;