**SQL ASSIGNMENT**

**Descripton :**

* The dataset contains details about the customers, products, sales and profit of a Superstore
* It basically includes 5 tables namely cust\_dimen (containing details about customer, their respective locations and the segment they belong to),prod\_dimen (containing product category and their subcategories), orders\_dimen (which provides us with the order date and order priority), shipping\_dimen (which gives us the shipping details), and market\_fact (which gives an overall idea about order quantity, sales value, discount, profit and shipping cost details etc)
* Different tables and their respective columns with datatype which is found using ‘describe’ feature in MySQL are provided below

1) CUST\_DIMEN:

Customer\_Name (TEXT): Name of the customer

Province (TEXT): Province of the customer

Region (TEXT): Region of the customer

Customer\_Segment (TEXT): Segment of the customer

Cust\_id (TEXT): Unique Customer ID

2)MARKET\_FACT:

Ord\_id (TEXT): Order ID

Prod\_id (TEXT): Product ID

Ship\_id (TEXT): Shipment ID

Cust\_id (TEXT): Customer ID

Sales (DOUBLE): Sales from the Item sold

Discount (DOUBLE): Discount on the Item sold

Order\_Quantity (INT): Order Quantity of the Item sold

Profit (DOUBLE): Profit from the Item sold

Shipping\_Cost (DOUBLE): Shipping Cost of the Item sold

Product\_Base\_Margin (DOUBLE): Product Base Margin on the Item sold

3)ORDERS\_DIMEN:

Order\_ID (INT): Order ID(seems irrelevant in this dataset)

Order\_Date (TEXT): Order Date

Order\_Priority (TEXT): Priority of the Order

Ord\_id (TEXT): Unique Order ID

4)PROD\_DIMEN:

Product\_Category (TEXT): Product Category

Product\_Sub\_Category (TEXT): Product Sub Category

Prod\_id (TEXT): Unique Product ID

5)SHIPPING\_DIMEN:

Order\_ID (INT): Order ID

Ship\_Mode (TEXT): Shipping Mode

Ship\_Date (TEXT): Shipping Date

Ship\_id (TEXT): Unique Shipment ID

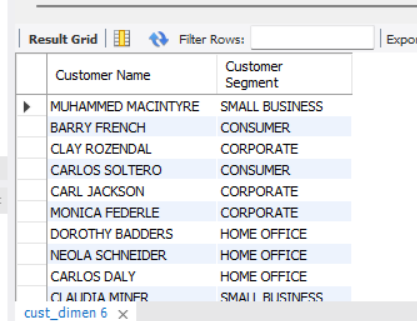
**PRIMARY & FOREIGN KEYS :**

1. cust\_dimen - Cust\_id is the Primary Key, there are no foreign keys
2. prod\_dimen - Prod\_id is the Primary Key, there are no foreign keys
3. orders\_dimen - Ord\_id is the Primary Key, No foreign key
4. shipping\_dimen - Shipping id as primary key and no foreign key.
5. market\_fact - Ord\_id, Prod\_id, Ship\_id and Cust\_id are the foreign keys. There are no Primary Keys.

**QUERIES :**

1. ***Write a query to display the Customer\_Name and Customer Segment using alias name “Customer Name", "Customer Segment" from table Cust\_dimen.***

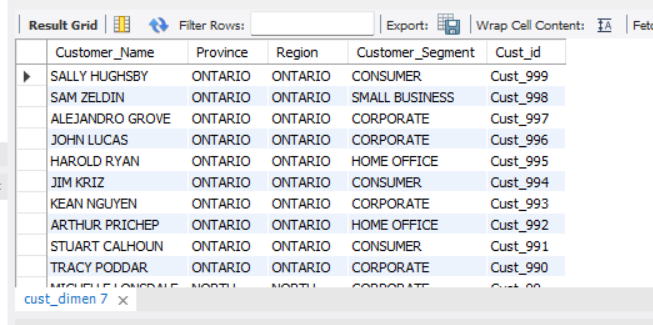
select Customer\_Name as 'Customer Name', Customer\_Segment as 'Customer Segment' from cust\_dimen ;



1. ***Write a query to find all the details of the customer from the table cust\_dimen order by desc***

select \* from cust\_dimen

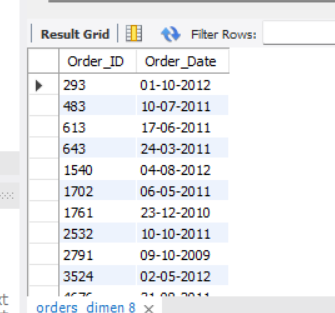
order by Cust\_id desc ;



1. ***Write a query to get the Order ID, Order date from table orders\_dimen where ‘Order Priority’ is high***.

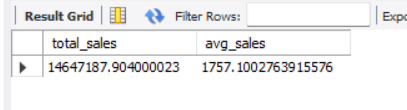
select Order\_ID,Order\_Date from orders\_dimen

where Order\_Priority = "High" ;



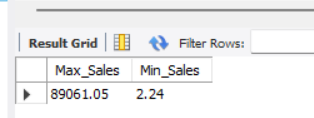
1. ***Find the total and the average sales (display total\_sales and avg\_sales)***

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



1. ***Write a query to get the maximum and minimum sales from maket\_fact table***

select max(Sales) Max\_Sales, min(Sales) Min\_Sales from market\_fact ;

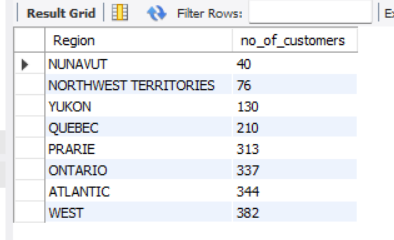


1. ***Display the number of customers in each region in decreasing order of no\_of\_customers. The result should contain columns Region, no\_of\_customers***

select Region,count(Cust\_id) as no\_of\_customers from cust\_dimen

group by Region

order by count(Cust\_id) ;

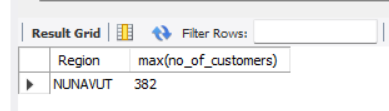


1. ***Find the region having maximum customers (display the region name and max(no\_of\_customers)***

select Region,max(no\_of\_customers) from

(select Region,count(Cust\_id) as no\_of\_customers from cust\_dimen

group by Region) as Maximum ;



1. ***Find all the customers from Atlantic region who have ever purchased ‘TABLES’ and the number of tables purchased (display the customer name, no\_of\_tables purchased)***

select c.Customer\_Name,sum(m.Order\_Quantity) as no\_of\_tables\_purchased

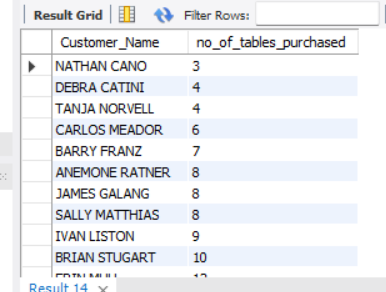
from market\_fact as m join cust\_dimen as c on m.Cust\_id=c.Cust\_id

join prod\_dimen as p on m.Prod\_id=p.Prod\_id

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

group by c.Customer\_Name

order by sum(m.Order\_Quantity) ;

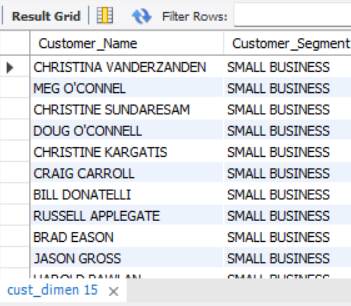


1. ***Find all the customers from Ontario province who own Small Business. (display the customer name, no of small business owners)***

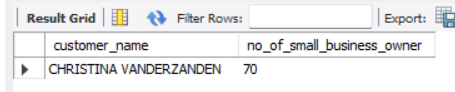
*# due to some confusion after reading the question I am providing two solutions here*

a] select Customer\_Name,Customer\_Segment from cust\_dimen

where Province= "ONTARIO" and Customer\_Segment= "SMALL BUSINESS" ;



b] select customer\_name,COUNT(\*) as no\_of\_small\_business\_owner from cust\_dimen where province='ontario' and customer\_segment='small business' ;



1. ***Find the number and id of products sold in decreasing order of products sold (display product id, no\_of\_products sold)***

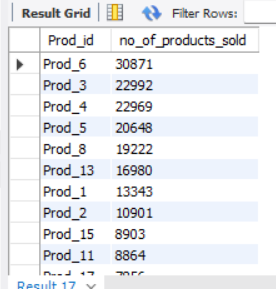
select p.Prod\_id,sum(m.Order\_quantity) as no\_of\_products\_sold

from market\_fact as m join prod\_dimen as p

on (m.Prod\_id=p.Prod\_id)

group by Prod\_id

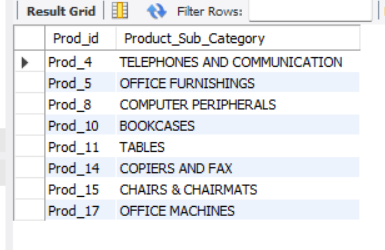
order by sum(m.Order\_quantity) desc ;



1. ***Display product Id and product sub category whose product category belongs to Furniture and Technology. The result should contain columns product id, product sub category***.

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

where Product\_Category in ("FURNITURE","TECHNOLOGY") ;



1. ***Display the product categories in descending order of profits (display the product category wise profits i.e. product\_category, profits***)

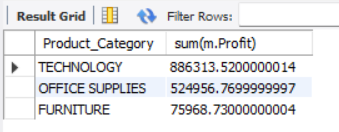
select p.Product\_Category,sum(m.Profit) from

market\_fact as m join prod\_dimen as p

on (m.Prod\_id=p.Prod\_id)

group by p.Product\_Category

order by sum(m.Profit) desc ;



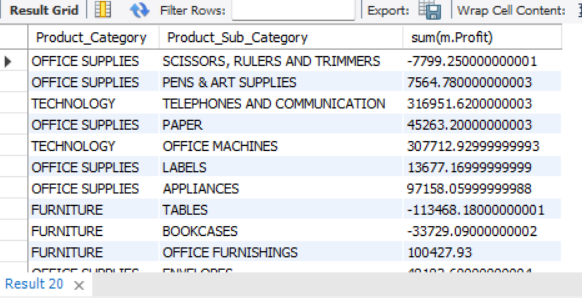
1. ***Display the product category, product sub-category and the profit within each subcategory in three columns.***

select p.Product\_Category,p.Product\_Sub\_Category,sum(m.Profit) from

market\_fact as m join prod\_dimen as p

on (m.Prod\_id=p.Prod\_id)

group by p.Product\_Sub\_Category ;

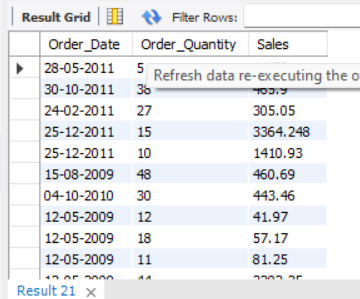


1. ***Display the order date, order quantity and the sales for the order.***

select o.Order\_Date,m.Order\_Quantity,m.Sales from

market\_fact as m join orders\_dimen as o

on (m.Ord\_id=o.Ord\_id);

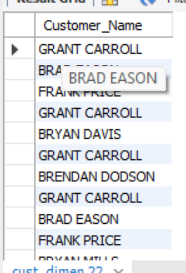


1. ***Display the names of the customers whose name contains the***
2. ***Second letter as ‘R’***
3. ***Fourth letter as ‘D’***

*# Due to some confusions after reading the question I am providing two solutions here*

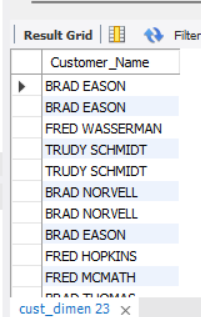
a] select Customer\_Name from cust\_dimen

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



b] select Customer\_Name from cust\_dimen

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



1. **Write a SQL query to to make a list with Cust\_Id, Sales, Customer Name and their region where sales are between 1000 and 5000.**

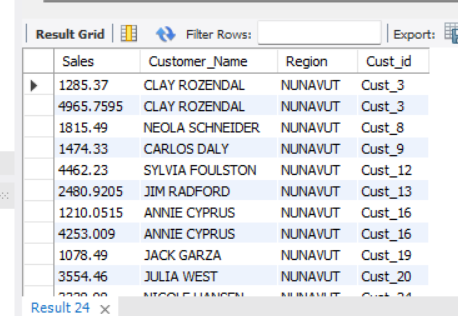
select m.Sales,c.Customer\_Name,c.Region,c.Cust\_id from

market\_fact as m join

cust\_dimen as c

on (m.Cust\_id=c.Cust\_id)

where sales between 1000 and 5000 ;



1. ***Write a SQL query to find the 3rd highest sales***

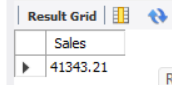
select Sales from (select Sales from market\_fact

order by Sales desc

limit 3) as High\_Sales

order by Sales asc

limit 1 ;



1. ***Where is the least profitable product subcategory shipped the most? For the least profitable product sub-category, display the region-wise no\_of\_shipments and the profit made in each region in decreasing order of profits (i.e. region, no\_of\_shipments, profit\_in\_each\_region)***

***→ Note: You can hardcode the name of the least profitable product subcategory***

select c.Region,count(m.Ship\_id) No\_of\_Shipments, sum(m.Profit) profit\_in\_each\_region

from market\_fact m

join cust\_dimen c on m.Cust\_id = c.Cust\_id

join prod\_dimen p on m.Prod\_id = p.Prod\_id

Where Product\_Sub\_Category =

( Select p.Product\_Sub\_Category from market\_fact m

join prod\_dimen p on m.Prod\_id = p.Prod\_id

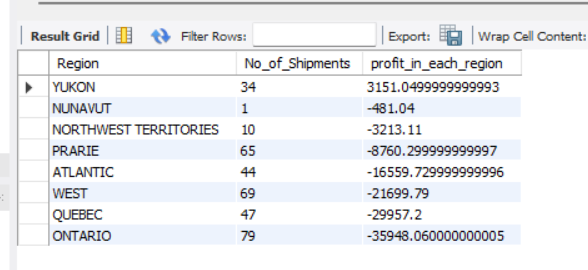
group by Product\_Sub\_Category

order by sum(m.Profit)

LIMIT 1)

group by c.Region

order by sum(m.Profit) desc;



The least profitable product subcategory is ‘TABLES’

By

**Alwin Mathew**