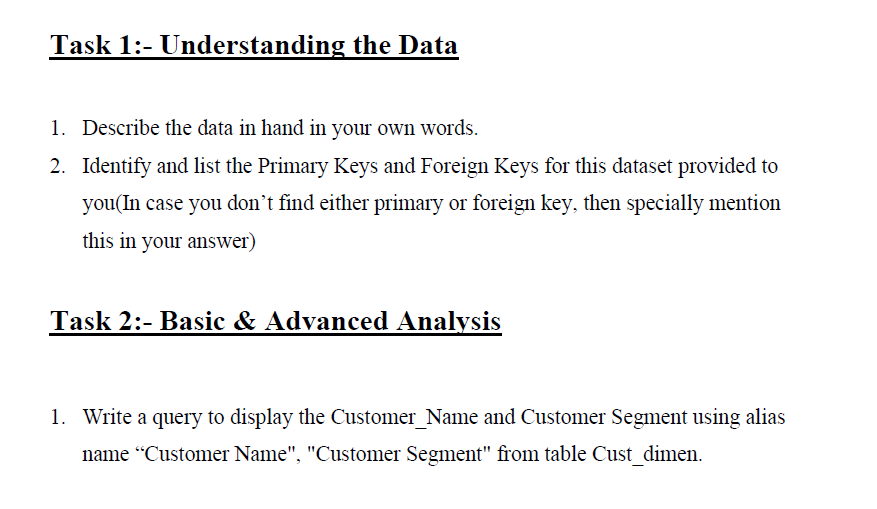
SQL ASSIGNMENT 2

Avinash A Godi





**---- Task 1 Understanding the Data ----**

**/\* 1. Describe the data in hand in your own words.**

This database contains Sales details of transaction of a superstores. The structure has 5 tables

**Table 1.** cust\_dimen (containing details about customer and their respective locations, contains 5 columns and 1833 rows)

**Column 1.** Customer\_Name (TEXT): Name of the customer

**Column 2.** Province (TEXT): Province of the customer

**Column 3.** Region (TEXT): Region of the customer

**Column 4.** Customer\_Segment (TEXT): Segment of the customer

**Column 5.** Cust\_id (TEXT): Unique Customer ID

**Table 2.** prod\_dimen (containing product category, product subcategories, and product id contains 3 columns and 18 rows)

**Column 1.** Product\_Category (TEXT): Product Category

**Column 2.** Product\_Sub\_Category (TEXT): Product Sub Category

**Column 3.** Prod\_id (TEXT): Unique Product ID

**Table 3.** orders\_dimen (containing order no, order date, order priority, and order id contains 4 columns and 5507 rows)

**Column 1.** Order\_ID (INT): Order ID

**Column 2.** Order\_Date (TEXT): Order Date

**Column 3.** Order\_Priority (TEXT): Priority of the Order

**Column 3.** Ord\_id (TEXT): Unique Order ID

**Table 4.** shipping\_dimen (containing ship date, order and shipping mode contains 4 columns and 7702 rows)

**Column 1.** Order\_ID (INT): Order ID

**Column 2.** Ship\_Mode (TEXT): Shipping Mode

**Column 3.** Ship\_Date (TEXT): Shipping Date

**Column 4.** Ship\_id (TEXT): Unique Shipment ID

**Table 5.** market\_fact (containing orderwise customerwise marketwise orderquantity, sales value, discount profit and shipping cost details contains 10 columns and 8400 rows).

**Column 1.** Ord\_id (TEXT): Order ID

**Column 2.** Prod\_id (TEXT): Prod ID

**Column 3.** Ship\_id (TEXT): Shipment ID

**Column 4.** Cust\_id (TEXT): Customer ID

**Column 5.** Sales (DOUBLE): Sales from the Item sold

**Column 6.** Discount (DOUBLE): Discount on the Item sold

**Column 7.** Order\_Quantity (INT): Order Quantity of the Item sold

**Column 8.** Profit (DOUBLE): Profit from the Item sold

**Column 9.** Shipping\_Cost (DOUBLE): Shipping Cost of the Item sold

**Column 10.** Product\_Base\_Margin (DOUBLE): Product Base Margin on the Item sold \*/

**/\* 2. 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 1.** cust\_dimen

Primary Key: Cust\_id

Foreign Key: NA

**Table 2.** prod\_dimen

Primary Key: Prod\_id, Product\_Sub\_Category

Foreign Key: NA

**Table 3.** orders\_dimen

Primary Key: Ord\_id

Foreign Key: NA

**Table 4.** shipping\_dimen

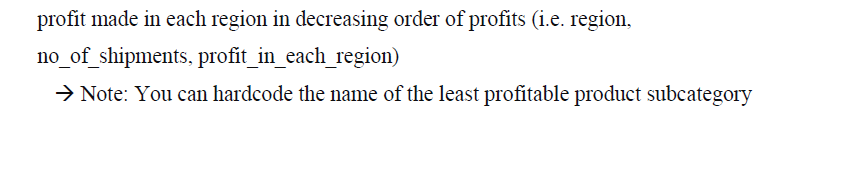
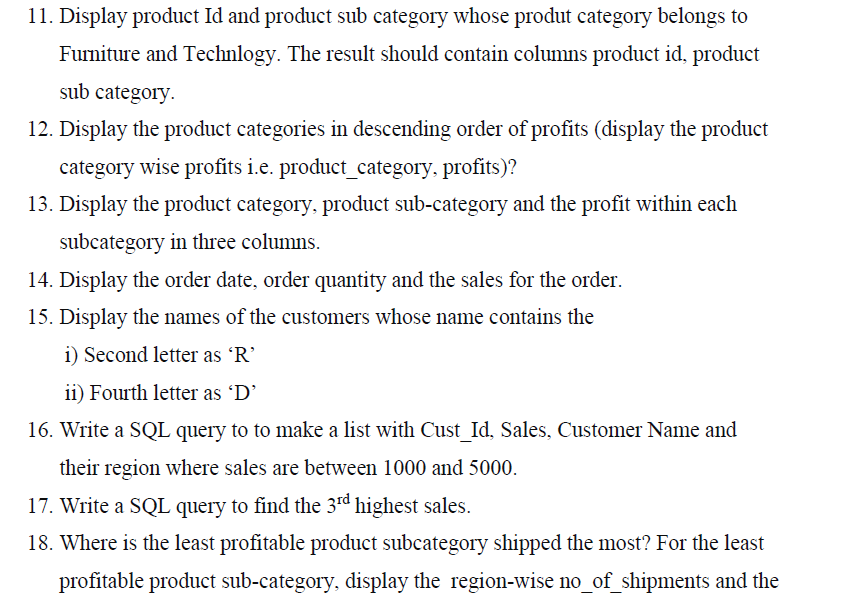
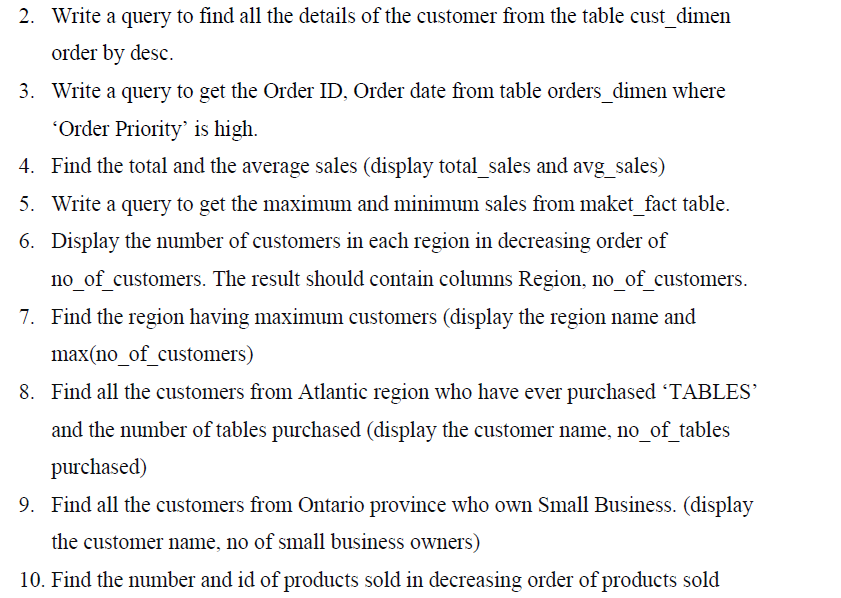
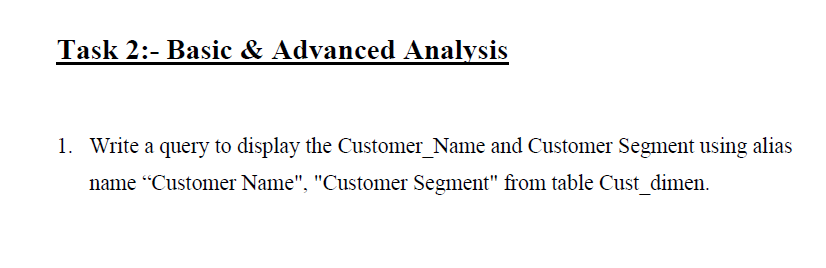
Primary Key: Ship\_id

Foreign Key: NA

**Table 5.** market\_fact

Primary Key: NA

Foreign Key: Ord\_id, Prod\_id, Ship\_id, Cust\_id \*/



**---- Task 2. Basic & Advanced Analysis ----**

**/\* 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 superstores.cust\_dimen;

**/\* 2. Write a query to find all the details of the customer from the table cust\_dimen order by desc. \*/**

select \* from superstores.cust\_dimen order by Customer\_Name desc;

**/\* 3. Write a query to get the Order ID, Order date from table orders\_dimen where ‘Order Priority’ is high. \*/**

select Ord\_id as 'Order ID', Order\_Date as 'Order Date', Order\_Priority from superstores.orders\_dimen where Order\_Priority like 'HIGH';

**/\* 4. Find the total and the average sales (display total\_sales and avg\_sales) \*/**

select round(sum(Sales),2) as 'Total Sales', round(avg(Sales),2) as 'Average Sales' from superstores.market\_fact;

**/\* 5. Write a query to get the maximum and minimum sales from maket\_fact table. \*/**

select max(Sales) as 'Maximum Sales', min(Sales) as 'Minimum Sales' from superstores.market\_fact;

**/\* 6. 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(\*) as 'No\_of\_customers' from superstores.cust\_dimen group by Region order by No\_of\_customers desc;

**/\* 7. Find the region having maximum customers (display the region name and max(no\_of\_customers). \*/**

select Region, count(\*) as 'No\_of\_customers' from superstores.cust\_dimen group by Region order by No\_of\_customers desc limit 1;

**/\* 8. 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 Customer\_Name, count(\*) as 'No\_of\_tables' from superstores.cust\_dimen c, superstores.market\_fact m, superstores.prod\_dimen p where c.Cust\_id=m.Cust\_id and m.Prod\_id=p.Prod\_id and Region like 'ATLANTIC' and Product\_Sub\_Category like 'TABLES' group by Customer\_Name;

**/\* 9. Find all the customers from Ontario province who own Small Business. (display the customer name, no of small business owners). \*/**

select Customer\_Name as 'Customer Name', Customer\_Segment as 'No. of Customer Segment', Province from superstores.cust\_dimen where Province like 'Ontario' and Customer\_Segment like 'small business';

**/\* 10. Find the number and id of products sold in decreasing order of products sold (display product id, no\_of\_products sold). \*/**

select Prod\_id, sum(Order\_Quantity) as "no\_of\_products sold" from superstores.market\_fact group by Prod\_id order by sum(Order\_Quantity) DESC;

**/\* 11. Display product Id and product sub category whose produt category belongs to Furniture and Technlogy. The result should contain columns product id, product sub category. \*/**

select Prod\_id as 'Product ID', Product\_Sub\_Category as 'Product Sub Category', Product\_Category as 'Product Category' from superstores.prod\_dimen where Product\_Category like 'TECHNOLOGY' or Product\_Category like 'FURNITURE';

**/\* 12. Display the product categories in descending order of profits (display the product category wise profits i.e. product\_category, profits)? \*/**

select Product\_Category as 'Product Category', round(sum(Profit),2) as 'Profits' from superstores.prod\_dimen p, superstores.market\_fact m where m.Prod\_id = p.Prod\_id group by Product\_Category order by sum(Profit) desc;

**/\* 13. Display the product category, product sub-category and the profit within each subcategory in three columns. \*/**

select Product\_Category as 'Product Category', Product\_Sub\_Category as 'Product Sub Category', round(sum(Profit),2) as 'Profits' from superstores.prod\_dimen p, superstores.market\_fact m where m.Prod\_id = p.Prod\_id group by Product\_Sub\_Category order by Product\_Category;

**/\* 14. Display the order date, order quantity and the sales for the order. \*/**

select Order\_Date, Order\_Quantity, Sales from superstores.orders\_dimen o, superstores.market\_fact m where m.Ord\_id = o.Ord\_id order by order\_quantity desc;

**/\* 15. Display the names of the customers whose name contains the \*/**

**/\* i) Second letter as ‘R’ \*/**

select Customer\_Name from superstores.cust\_dimen where Customer\_Name like '\_r%';

**/\* ii) Fourth letter as ‘D’ \*/**

select Customer\_Name from superstores.cust\_dimen where Customer\_Name like '\_\_\_d%';

**/\* 16. 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 c.Cust\_id, Sales, Customer\_Name as 'Customer Name', Region from superstores.cust\_dimen c, superstores.market\_fact m where m.Cust\_id = c.Cust\_id and Sales between 1000 and 5000;

**/\* 17. Write a SQL query to find the 3rd highest sales. \*/**

select Sales as 'Third Highest Salary' from superstores.market\_fact order by Sales desc limit 2,1;

**/\* 18. 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 as "Region",count(m.Ship\_id) as "No of Shipments", round(sum(m.Profit),2) as "Profit in each region"

from superstores.market\_fact m

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

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

Where Product\_Sub\_Category = (Select p.Product\_Sub\_Category from superstores.market\_fact m join superstores.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);