SQL ASSIGNMENT 2

Task – 1: Understanding the Data

Q.1 Describe the data in your own words.

Ans:

1. CUST_DIMEN: Details of all the Customers

   Customer_Name (string): Name of the customer

   Province (string): Province of the customer

   Region (string): Region of the customer

   Customer_Segment (string): Segment of the customer

   Cust_ID (string): Unique customer ID

2. MARKET_FACT: Details of every order item sold.

   Ord_ID (string): Order ID
   Prod_ID (string): Prod ID
   Ship_ID (string): Shipment ID
   Cust_ID (string): Customer ID
   Sales (double): Sales from the Item sold
   Discount (double): Discount on the Item sold
   Order_Quality (int): Order Quantity of the Item sold
   Profit (double): Profit from the sold Item
   Shipping_cost (double): Shipping cost of the Item sold
   Product_Based_Margin (double): Product Based Margin on the Item sold

3. ORDERS_DIMEN: Details of every order placed.

   Order_ID (int): Order ID

   Order_Date (string): Order Date

   Order_Priority (string): Priority of the order

   Ord_id (string): Unique order ID

4. SHIPPING_DIMEN: Details of Shipping of Orders
   Order_ID (int): Order ID
   Ship_Mode (string): Shipping mode
   Ship_Date (string): Shipping Date
   Ship_ID (string): Unique Shipment ID

Q.2 Identify and list the Primary Keys and Foreign Keys for this dataset provided to you.

Ans:

1) **PRIMARY KEY**

CUST_DIMEN: CUST_ID

MARKET_FACT: NO PRIMARY KEY

ORDERS_DIMEN: ORD_ID

PROD_DIMEN: PROD_ID

SHIPPING_DIMEN: SHIP_ID

2) **FOREIGN KRY**

MARKET_FACT: ORD_ID, PROD_ID, SHIP_ID, CUST_ID

ORDERS_DIMEN: ORDER_ID

SHIPPING_DIMEN: ORDER_ID

TASK 2: Basic & Advanced Analysis

```
CREATE DATABASE SUPERSTORE;        #importing all CSV files
USE SUPERSTORE;
```

Q.1 Write a query to display the Customer_Name and Customer Segment using alias name “Customer Name", "Customer Segment" from table Cust_dimen.

Ans:
```
SELECT CUSTOMER_NAME "CUSTOMER NAME", CUSTOMER_SEGMENT "CUSTOMER SEGMENT" FROM CUST_DIMEN;
```

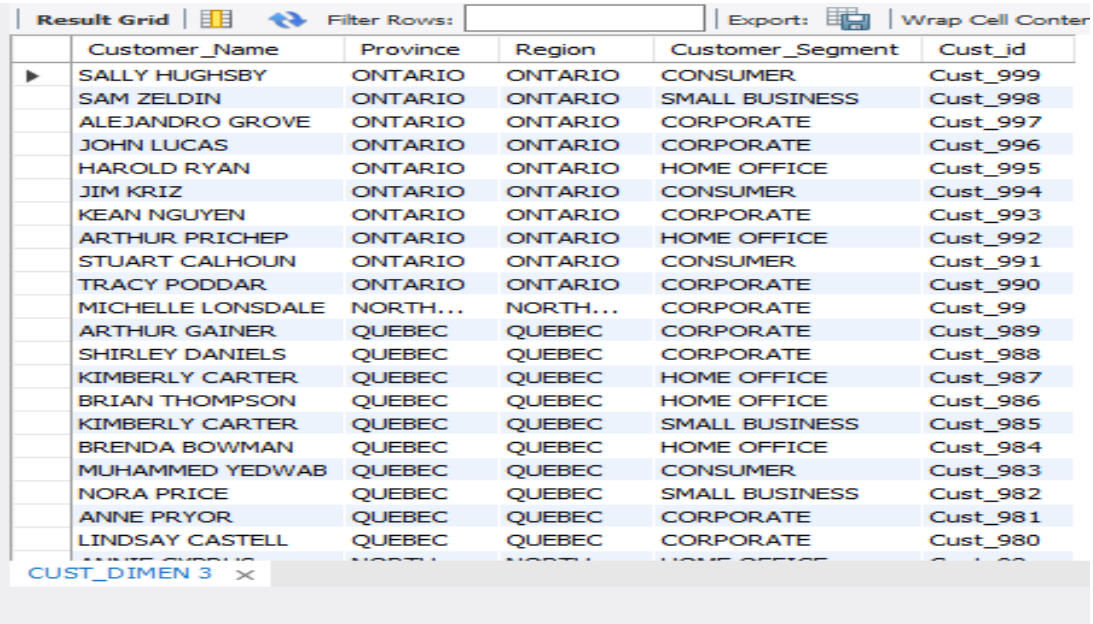
Result Grid | Filter Rows:

CUSTOMER NAME	CUSTOMER SEGMENT
MUHAMMED MACINTYRE	SMALL BUSINESS
BARRY FRENCH	CONSUMER
CLAY ROZENDAL	CORPORATE
CARLOS SOLTERO	CONSUMER
CARL JACKSON	CORPORATE
MONICA FEDERLE	CORPORATE
DOROTHY BADDERS	HOME OFFICE
NEOLA SCHNEIDER	HOME OFFICE
CARLOS DALY	HOME OFFICE
CLAUDIA MINER	SMALL BUSINESS
ALLEN ROSENBLATT	SMALL BUSINESS
SYLVIA FOULSTON	HOME OFFICE
JIM RADFORD	CORPORATE
CARL LUDWIG	CORPORATE
DON MILLER	HOME OFFICE
ANNIE CYPRUS	HOME OFFICE
GRANT CARROLL	SMALL BUSINESS
ALAN BARNES	CORPORATE
JACK GARZA	CORPORATE
JULIA WEST	CORPORATE
EUGENE BARCHAS	CORPORATE

UST_DIMEN 1

Q.2 Write a query to find all the details of the customer from the table cust_dimen order by desc.

Ans: SELECT * FROM CUST_DIMEN ORDER BY CUST_ID DESC;



Result Grid | Filter Rows: | Export: | Wrap Cell Conter

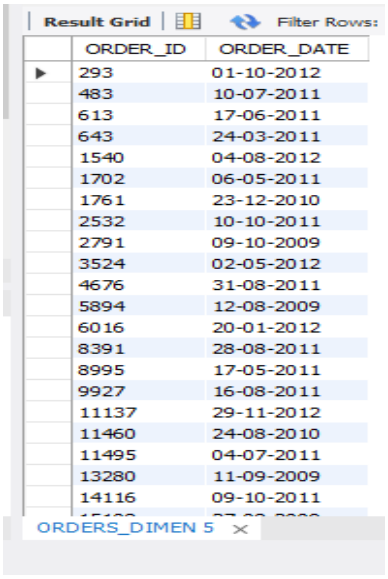
Customer_Name	Province	Region	Customer_Segment	Cust_id
SALLY HUGHSBY	ONTARIO	ONTARIO	CONSUMER	Cust_999
SAM ZELDIN	ONTARIO	ONTARIO	SMALL BUSINESS	Cust_998
ALEJANDRO GROVE	ONTARIO	ONTARIO	CORPORATE	Cust_997
JOHN LUCAS	ONTARIO	ONTARIO	CORPORATE	Cust_996
HAROLD RYAN	ONTARIO	ONTARIO	HOME OFFICE	Cust_995
JIM KRIZ	ONTARIO	ONTARIO	CONSUMER	Cust_994
KEAN NGUYEN	ONTARIO	ONTARIO	CORPORATE	Cust_993
ARTHUR PRICHEP	ONTARIO	ONTARIO	HOME OFFICE	Cust_992
STUART CALHOUN	ONTARIO	ONTARIO	CONSUMER	Cust_991
TRACY PODDAR	ONTARIO	ONTARIO	CORPORATE	Cust_990
MICHELLE LONSDALE	NORTH...	NORTH...	CORPORATE	Cust_99
ARTHUR GAINER	QUEBEC	QUEBEC	CORPORATE	Cust_989
SHIRLEY DANIELS	QUEBEC	QUEBEC	CORPORATE	Cust_988
KIMBERLY CARTER	QUEBEC	QUEBEC	HOME OFFICE	Cust_987
BRIAN THOMPSON	QUEBEC	QUEBEC	HOME OFFICE	Cust_986
KIMBERLY CARTER	QUEBEC	QUEBEC	SMALL BUSINESS	Cust_985
BRENDA BOWMAN	QUEBEC	QUEBEC	HOME OFFICE	Cust_984
MUHAMMED YEDWAB	QUEBEC	QUEBEC	CONSUMER	Cust_983
NORA PRICE	QUEBEC	QUEBEC	SMALL BUSINESS	Cust_982
ANNE PRYOR	QUEBEC	QUEBEC	CORPORATE	Cust_981
LINDSAY CASTELL	QUEBEC	QUEBEC	CORPORATE	Cust_980

CUST_DIMEN 3

Q.3 Write a query to get the Order ID, Order date from table orders_dimen where 'Order Priority' is high

Ans: SELECT ORDER_ID, ORDER_DATE FROM ORDERS_DIMEN

WHERE ORDER_PRIORITY = "HIGH";



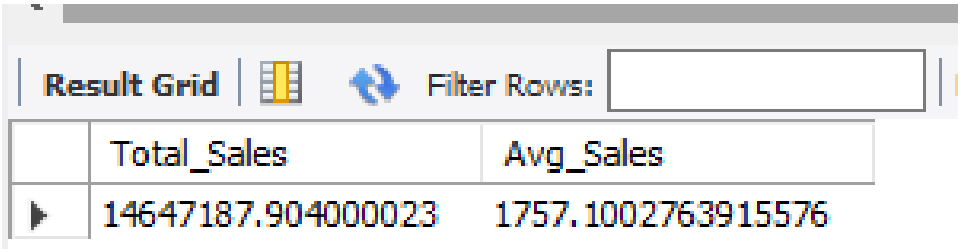
Result Grid | Filter Rows:

ORDER_ID	ORDER_DATE
293	01-10-2012
483	10-07-2011
613	17-06-2011
643	24-03-2011
1540	04-08-2012
1702	06-05-2011
1761	23-12-2010
2532	10-10-2011
2791	09-10-2009
3524	02-05-2012
4676	31-08-2011
5894	12-08-2009
6016	20-01-2012
8391	28-08-2011
8995	17-05-2011
9927	16-08-2011
11137	29-11-2012
11460	24-08-2010
11495	04-07-2011
13280	11-09-2009
14116	09-10-2011

ORDERS_DIMEN 5

Q.4 Find the total and the average sales (display total_sales and avg_sales)

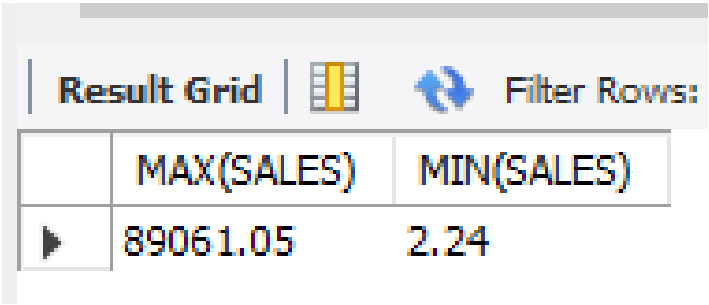
Ans: SELECT SUM(SALES) "Total_Sales", AVG(SALES) "Avg_Sales" FROM MARKET_FACT;



Q.5 Write a query to get the maximum and minimum sales from market_fact table.

Ans: SELECT MAX(SALES), MIN(SALES)

FROM MARKET_FACT;



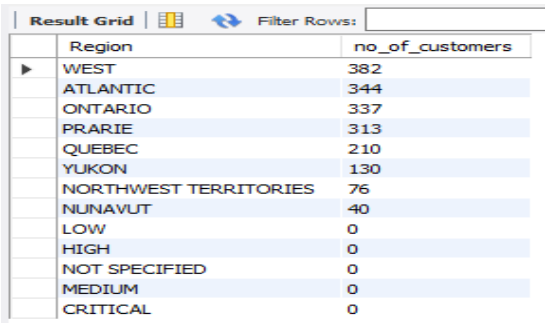
Q.6 Display the number of customers in each region in decreasing order of no_of_customers. The result should contain columns Region, no_of_customers.

Ans: SELECT REGION as Region, COUNT(CUST_ID) "no_of_customers"

FROM CUST_DIMEN

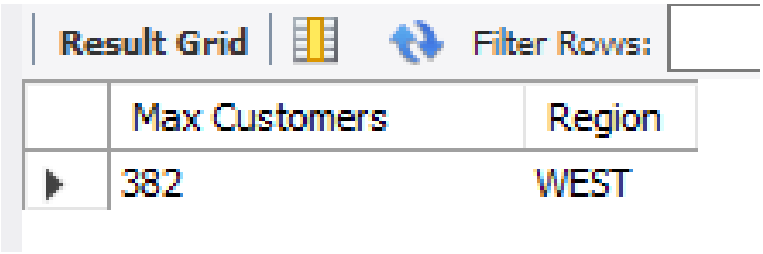
GROUP BY Region

ORDER BY COUNT(CUST_ID) DESC;



Q.7 Find the region having maximum customers (display the region name and max(no_of_customers)

Ans:
```
SELECT MAX(no_of_cust) "Max Customers", Region
FROM (SELECT REGION, COUNT(CUST_ID) "no_of_cust"
FROM CUST_DIMEN
GROUP BY REGION
ORDER BY COUNT(CUST_ID) DESC) AS MAX_CUST_REGION;
```



Q.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)

Ans:
```
SELECT Customer_Name, COUNT(p.Prod_id) AS
'no_of_tables_purchase'
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 c.Region='ATLANTIC' AND p.Product_Sub_Category='TABLES'
GROUP BY c.Customer_Name ORDER BY COUNT(c.Cust_id) DESC;
```

Customer_Name	no_of_tables_purchase
MUHAMMED YEDWAB	2
PATRICK JONES	2
BARRY FRANZ	1
BECKY MARTIN	1
BEN PETERMAN	1
BOBBY TRAFTON	1
BRADLEY TALBOTT	1
BRIAN STUGART	1
CARLOS MEADOR	1
CARLOS SOLTERO	1
CATHY ARMSTRONG	1
CHRISTINA DEMOSS	1
CHRISTY BRITTAIN	1
CHUCK CLARK	1
CHUCK SACHS	1
CRAIG YEDWAB	1
DAVID FLASHING	1
DEBRA CATINI	1
DENISE LEINENBACH	1
ELENI MCCRARY	1
ERICA BERN	1

Result 25

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

Ans:

```
SELECT Customer_Name,
COUNT(Customer_Name) AS 'No of small business owners'
FROM cust_dimen
WHERE Customer_Segment='SMALL BUSINESS' AND Region='ONTARIO'
GROUP BY Customer_name ;
```

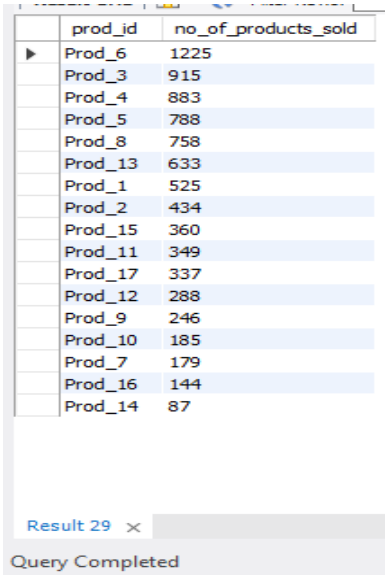
Customer_Name	No of small business owners
CHRISTINA VANDERZANDEN	1
MEG O'CONNEL	1
CHRISTINE SUNDARESAM	1
DOUG O'CONNELL	1
CHRISTINE KARGATIS	1
CRAIG CARROLL	1
BILL DONATELLI	1
RUSSELL APPLEGATE	1
BRAD EASON	1
JASON GROSS	1
HAROLD PAWLAN	1
ADAM BELLAVANCE	1
PAULINE CHAND	1
MICHAEL NGUYEN	1
HENRY MACALLISTER	1
DUANE HUFFMAN	1
JEREMY ELLISON	1
FRANK CARLISLE	1
CHRISTINE PHAN	1
ALAN HWANG	1
JESUS OCAMPO	1

Result 27

Query Completed

Q.10 Find the number and id of products sold in decreasing order of products sold (display product id, no_of_products sold)

Ans:
```
SELECT prod_id ,
COUNT(prod_id) AS no_of_products_sold
FROM MARKET_FACT
GROUP BY Prod_id
ORDER BY COUNT(prod_id) DESC;
```



prod_id	no_of_products_sold
Prod_6	1225
Prod_3	915
Prod_4	883
Prod_5	788
Prod_8	758
Prod_13	633
Prod_1	525
Prod_2	434
Prod_15	360
Prod_11	349
Prod_17	337
Prod_12	288
Prod_9	246
Prod_10	185
Prod_7	179
Prod_16	144
Prod_14	87

Result 29 ×

Query Completed

Q.11 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.

Ans:
```
SELECT prod_id "product id",
product_sub_category AS "product sub category"
FROM PROD_DIMEN
WHERE PRODUCT_CATEGORY = "TECHNOLOGY" OR
PRODUCT_CATEGORY = "FURNITURE";
```

	product id	product sub category
▶	Prod_4	TELEPHONES AND COMMUNICATION
	Prod_5	OFFICE FURNISHINGS
	Prod_8	COMPUTER PERIPHERALS
	Prod_10	BOOKCASES
	Prod_11	TABLES
	Prod_14	COPIERS AND FAX
	Prod_15	CHAIRS & CHAIRMATS
	Prod_17	OFFICE MACHINES

PROD_DIMEN 37 ×

Query Completed

Q.12 Display the product categories in descending order of profits (display the product category wise profits i.e. product_category, profits)?

Ans:
```
SELECT p.Product_Category, COUNT(m.Profit) AS profits
FROM market_fact m
JOIN prod_dimen p ON p.Prod_id=m.Prod_id
GROUP BY p.Product_Category
ORDER BY COUNT(m.Profit) DESC;
```

	Product_Category	profits
▶	OFFICE SUPPLIES	4589
	TECHNOLOGY	2065
	FURNITURE	1682

Result 40 ×

Query Completed

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

Ans:
```
SELECT PROD.PRODUCT_CATEGORY,
PROD.PRODUCT_SUB_CATEGORY,
COUNT(MARKET.PROFIT) AS PROFIT
FROM MARKET_FACT AS MARKET
JOIN PROD_DIMEN AS PROD ON
```

PROD.PROD_ID = MARKET.PROD_ID

GROUP BY PROD.PRODUCT_SUB_CATEGORY;

PRODUCT_CATEGORY	PRODUCT_SUB_CATEGORY	PROFIT
OFFICE SUPPLIES	SCISSORS, RULERS AND TRIMMERS	144
OFFICE SUPPLIES	PENS & ART SUPPLIES	633
TECHNOLOGY	TELEPHONES AND COMMUNICATION	883
OFFICE SUPPLIES	PAPER	1225
TECHNOLOGY	OFFICE MACHINES	337
OFFICE SUPPLIES	LABELS	288
OFFICE SUPPLIES	APPLIANCES	434
FURNITURE	TABLES	349
FURNITURE	BOOKCASES	185
FURNITURE	OFFICE FURNISHINGS	788
OFFICE SUPPLIES	ENVELOPES	246
FURNITURE	CHAIRS & CHAIRMATS	360
OFFICE SUPPLIES	RUBBER BANDS	179
OFFICE SUPPLIES	BINDERS AND BINDER ACCESSORIES	915
OFFICE SUPPLIES	STORAGE & ORGANIZATION	525
TECHNOLOGY	COMPUTER PERIPHERALS	758
TECHNOLOGY	COPIERS AND FAX	87

Result 48

Query Completed

Q.14 Display the order date, order quantity and the sales for the order.

Ans: SELECT ord.Order_date as "Order Date",

mark.Order_Quantity as "Order Quantity",

mark.sales "Sales"

FROM superstore.market_fact AS mark JOIN superstore.orders_dimen AS ord

ON ord.Ord_id = mark.Ord_id;

Result Grid | Filter Rows:

Order Date	Order Quantity	Sales
28-05-2011	5	14.76
30-10-2011	38	465.9
24-02-2011	27	305.05
25-12-2011	15	3364.248
25-12-2011	10	1410.93
15-08-2009	48	460.69
04-10-2010	30	443.46
12-05-2009	12	41.97
12-05-2009	18	57.17
12-05-2009	11	81.25
12-05-2009	44	3202.25
12-05-2009	10	35.64
12-02-2009	13	197.61
04-12-2010	22	38.26
19-01-2012	13	109.58
28-05-2009	28	1062.69
12-02-2009	38	3594.7435
21-01-2009	33	139.98
08-04-2009	27	129.1
12-02-2009	22	68.92
29-04-2011	12	76.16

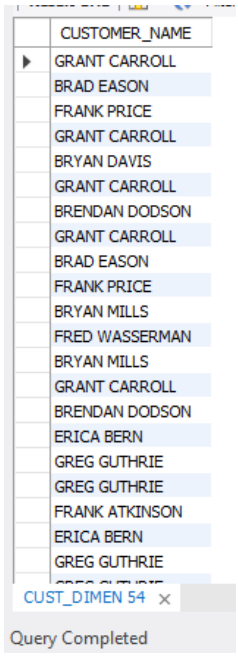
Result 53

Query Completed

Q.15 Display the names of the customers whose name contains the

i) Second letter as 'R'

Ans:
```
SELECT CUSTOMER_NAME FROM CUST_DIMEN
WHERE CUSTOMER_NAME LIKE "_R%";
```



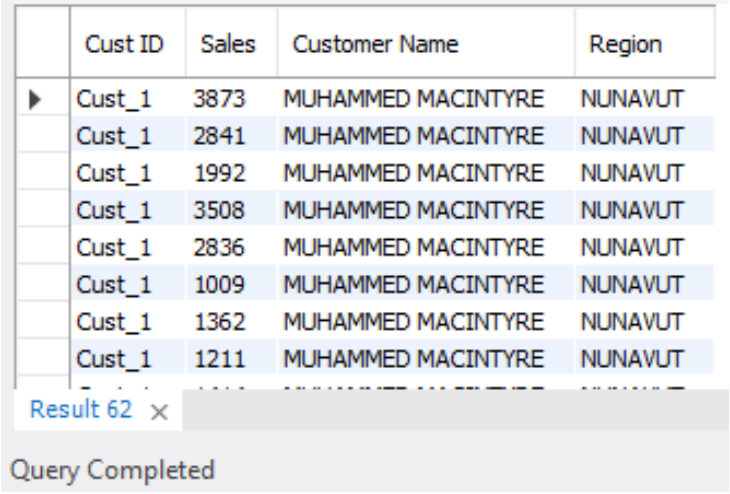
CUSTOMER_NAME
GRANT CARROLL
BRAD EASON
FRANK PRICE
GRANT CARROLL
BRYAN DAVIS
GRANT CARROLL
BRENDAN DODSON
GRANT CARROLL
BRAD EASON
FRANK PRICE
BRYAN MILLS
FRED WASSERMAN
BRYAN MILLS
GRANT CARROLL
BRENDAN DODSON
ERICA BERN
GREG GUTHRIE
GREG GUTHRIE
FRANK ATKINSON
ERICA BERN
GREG GUTHRIE

ii)) Fourth letter as 'D'

Ans:
```
SELECT CUSTOMER_NAME FROM CUST_DIMEN
WHERE CUSTOMER_NAME LIKE "___D%";
```

Q.16 Write a SQL query to make a list with Cust_Id, Sales, Customer Name and their region where sales are between 1000 and 5000.

Ans:
```
SELECT C.CUST_ID "Cust ID", ROUND(MARK.SALES) "Sales",
C.CUSTOMER_NAME "Customer Name", C.REGION "Region"
FROM CUST_DIMEN AS C
JOIN MARKET_FACT AS MARK
WHERE SALES BETWEEN 1000 AND 5000;
```

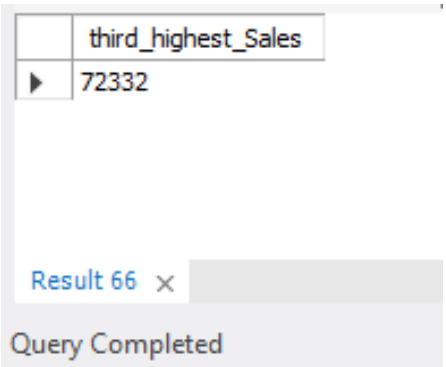


Cust ID	Sales	Customer Name	Region
Cust_1	3873	MUHAMMED MACINTYRE	NUNAVUT
Cust_1	2841	MUHAMMED MACINTYRE	NUNAVUT
Cust_1	1992	MUHAMMED MACINTYRE	NUNAVUT
Cust_1	3508	MUHAMMED MACINTYRE	NUNAVUT
Cust_1	2836	MUHAMMED MACINTYRE	NUNAVUT
Cust_1	1009	MUHAMMED MACINTYRE	NUNAVUT
Cust_1	1362	MUHAMMED MACINTYRE	NUNAVUT
Cust_1	1211	MUHAMMED MACINTYRE	NUNAVUT

Q.17 Write a SQL query to find the 3rd highest sales.

Ans:

```
SELECT MIN(sales) AS "3rd highest sales"
FROM (SELECT ROUND(SUM(Sales)) AS sales
FROM market_fact
GROUP BY Cust_id
ORDER BY SUM(Sales) DESC LIMIT 3)
AS Tab1;
```



third_highest_Sales
72332

Q.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

Ans:

```
SELECT c.Region, COUNT(s.Ship_id) AS no_of_shipment,
ROUND(SUM(m.Profit)) AS profit
FROM market_fact  m
JOIN cust_dimen c ON c.Cust_id=m.Cust_id
JOIN prod_dimen p ON p.Prod_id= m.Prod_id
JOIN shipping_dimen s ON s.Ship_id = m.Ship_id
WHERE p.Product_Sub_Category='TABLES'
GROUP BY c.Region
ORDER BY ROUND(SUM(m.Profit)) DESC;
```

Region	no_of_shipment	profit
YUKON	34	3151
NUNAVUT	1	-481
NORTHWEST TERRITORIES	10	-3213
PRARIE	65	-8760
ATLANTIC	44	-16560
WEST	69	-21700
QUEBEC	47	-29957
ONTARIO	79	-35948

Result 67 ×

Query Completed

- **GANESH DEVARE**