

Joshua Merren

Professor Lockwood

DAD-220 Intro to Struct Database Env

5 February 2024

### 5-3 Major Activity: Analyzing Queries

1. **Write SQL commands** that capture specific, usable data that can be used in your analysis.
2. **Analyze the results of queries** to identify specific information that can be presented in your summary.

#### A. Sales by region:

- i. Analyze sales data by state to determine where the company has the largest customer base.

```
mysql> SELECT
->   @row_num := @row_num + 1 AS state_rank,
->   State,
->   total_unique_customers
-> FROM (
->   SELECT
->     c.State,
->     COUNT(DISTINCT o.CustomerID) AS total_unique_customers
->   FROM Orders AS o
->   INNER JOIN Customers AS c ON o.CustomerID = c.CustomerID
->   GROUP BY c.State
->   ) AS subquery, (SELECT @row_num := 0) AS t
-> ORDER BY total_unique_customers DESC;
```

state_rank	State	total_unique_customers
1	Massachusetts	982
2	Arkansas	854
3	West Virginia	843
4	Oregon	842
5	Alabama	838
6	Idaho	838
7	Mississippi	834
8	Tennessee	831
9	Connecticut	830
10	Delaware	824
11	Wisconsin	822

This screenshot shows that Massachusetts has the largest consumer base. A select statement was used to find the largest consumer base and to have it ordered by largest to smallest which is South Carolina.

- ii. Analyze the data to determine the top three products sold in the United States.

```
mysql> SELECT COUNT(*) AS total, Orders.SKU, Orders.Description AS Description
-> FROM Orders
-> GROUP BY SKU, Description
-> ORDER BY total DESC
-> LIMIT 3;
```

total	SKU	Description
8385	BAS-48-1 C	Basic Switch 10/100/1000 BaseT 48 port
6186	ENT-48-40F	Enterprise Switch 40GigE SFP+ 48 port
4328	ENT-48-10F	Enterprise Switch 10GigE SFP+ 48 port

3 rows in set (0.56 sec)

This screenshot shows that the top three sellers are Basic Switch 10/100/1000, Enterprise Switch 40Gig, and Enterprise Switch 10Gig. I found this by using a Select query to get the top 3 orders by SKU.

- iii. Analyze the data to determine the top three products sold in the southeastern region of the United States.

```
mysql> SELECT COUNT(*) AS PRODUCT_SALES_NUMBER, Orders.SKU AS PRODUCT_SKU, MAX(Orders.Description) AS PRODUCT_DESCRIPTION
-> FROM Orders
-> INNER JOIN Customers ON Orders.CustomerID = Customers.CustomerID
-> WHERE UPPER(State) IN ("VIRGINIA", "NORTH CAROLINA", "SOUTH CAROLINA", "GEORGIA")
-> GROUP BY PRODUCT_SKU
-> ORDER BY PRODUCT_SALES_NUMBER DESC
-> LIMIT 10;
```

PRODUCT_SALES_NUMBER	PRODUCT_SKU	PRODUCT_DESCRIPTION
	S-48-1 C	Basic Switch 10/100/1000 BaseT 48 port
	T-48-40F	Enterprise Switch 40GigE SFP+ 48 port
	S-08-1 C	Basic Switch 10/100/1000 BaseT 8 port
256	ADV-48-10F	Advanced Switch 10 GigE Copper/Fiber 44 port copper 4 port fiber
	-48-10F	Enterprise Switch 10GigE SFP+ 48 port
	-24-10C	Advanced Switch 10GigE Copper 24 port
	-24-10F	Enterprise Switch 10GigE SFP+ 24 Port
	T-24-40F	Enterprise Switch 40GigE SFP+ 24 port
	S-24-1 C	Basic Switch 10/100/1000 BaseT 24 port

9 rows in set (0.11 sec)

This screenshot shows the top three products sold in the United States are Basic Switch 10/100/1000 BaseT 48 port, Enterprise Switch 40GigE SEP+ 48 port, and Basic Switch 10/100/1000 BaseT 8 port. I found this by using a select query for the top 10 SKUs by count that were returned.

## B. Returns by region:

- i. Analyze the data to determine the top three products returned in the United States.

```
ll_group_by
mysql> SELECT COUNT(*) AS returned_amount, Orders.SKU AS product_sku, MAX(Orders.Description) AS product_description
-> FROM Orders
-> INNER JOIN RMA ON Orders.orderID = RMA.OrderID
-> WHERE UPPER(status) = 'COMPLETE'
-> GROUP BY product_sku
-> ORDER BY returned_amount DESC
-> LIMIT 3;
```

returned_amount	product_sku	product_description
7405	BAS-48-1 C	Basic Switch 10/100/1000 BaseT 48 port
5470	ENT-48-40F	Enterprise Switch 40GigE SFP+ 48 port
3847	ENT-48-10F	Enterprise Switch 10GigE SFP+ 48 port

3 rows in set (0.96 sec)

```
mysql>
```

This screenshot shows the top three products returned in the United States. Basic Switch 10/100/1000 BaseT 48 port, Enterprise Switch 48GigE SFP+ 48 port, and Enterprise Switch 10GigE SFP+ 48 port are the top 3 most returned products in the United States. This was found by using a select query on the top SKUs by counting for the most returned.

- ii. Analyze the data to determine the top three products returned in the northwestern region of the United States.

- Northwestern states to include in your analysis: Washington, Oregon, Idaho, and Montana

```
mysql> SELECT COUNT(*) AS returned_amount, Orders.SKU AS product_SKU, MAX(Orders.Description) AS product_description
-> FROM Orders
-> INNER JOIN RMA ON Orders.OrderID = RMA.OrderID
-> INNER JOIN Customers ON Orders.CustomerID = Customers.CustomerID
-> WHERE UPPER(State) IN ('WASHINGTON', 'OREGON', 'IDAHO', 'MONTANA') AND UPPER(Status) = 'COMPLETE'
-> GROUP BY Orders.SKU
-> ORDER BY returned_amount DESC
-> LIMIT 3;
```

returned_amount	product_SKU	product_description
628	BAS-48-1 C	Basic Switch 10/100/1000 BaseT 48 port
473	ENT-48-40F	Enterprise Switch 40GigE SFP+ 48 port
344	ENT-24-10F	Enterprise Switch 10GigE SFP+ 24 Port

3 rows in set (0.21 sec)

This screenshot shows the top three items returned in the northwest United States. They are BAS-48-1 C, ENT-48-40F, and ENT-24-10F. This was determined by using a select query for the top three SKUs by count that were returned and narrowed down to just these states.

3. **Write a report** to the Quantigration product manager that explains your findings in a way nontechnical stakeholders can digest and use.

**A.** This report should include an effective summary of the analysis of the captured data.

**i.** Sales data by region: Provide a well-written summary of your analysis on Part A.

In analyzing sales data by region, I utilized an inner join between the Customers and Orders tables, aiming to determine the customer base distribution across states. My findings revealed that Massachusetts holds the largest customer base, boasting 982 orders. Arkansas followed this with 854 orders, and West Virginia had 843 orders. Moving on to the top-selling products in the United States, the analysis identified Basic Switch 48 port, Enterprise Switch 48 port, and Enterprise Switch 10Gig 48 port as the top three products sold.

For the Southeastern region, the top three products sold were Basic Switch 48 port, Enterprise Switch 48 port, and Basic Switch 8 port. These findings provide valuable insights into regional sales patterns and product preferences. The queries involved using inner joins, count commands, and where clauses to filter specific states, ensuring a comprehensive and targeted sales data analysis.

**ii.** Returns data by region: Provide a well-written summary of your analysis of Part B.

Analyzing returns data by region involved a join between the Orders and RMA tables, focusing on completed orders. The top three returned products in the United States were identified as Basic Switch BaseT 48 port, Enterprise Switch 48 port, and Enterprise Switch 48 port. The top three returned products for the Northwestern region mirrored the national trends, featuring Basic Switch BaseT 48 port, Enterprise Switch 48 port, and Enterprise Switch 24 port.

Upon comparing the sales and returns data, it can be observed that the top three returned products closely align with the top-selling items. For instance, the BaseT 48 port had 7545 units sold and 645 units returned. This suggests a potential area for further product quality and customer satisfaction investigation. The queries employed in this analysis included joins, filters for completed orders, and limitations to capture the top three returned products in descending order. These results provide valuable information that can be leveraged to improve the quality of our products and enhance the overall customer experience.