Analysis on supermart_DB

Download and import the supertmart_DB into your pgadmin

1. Sales Performance by Region

- *Question:* What is the total sales amount by region?
- **Hint***:* Use the `sales` and `customer` tables to join the data based on `customer_id`.

2. Top-Selling Products

- *Question:* Which products generated the most sales?
- **Hint***:* Join the `sales` and `product` tables on `product_id` and sum the `sales` column.

3. Discount Impact on Profit

- *Question:* How does the discount affect profit?
- **Hint***:* Query the `sales` table to compare profit and discount.

4. Sales by Customer Segment

- *Question:* How much sales does each customer segment contribute?
- **Hint***:* Use the `sales` and `customer` tables, grouping the data by `segment`.

5. Product Category Sales

- *Question:* What are the total sales for each product category?
- **Hint***:* Join the `sales` and `product` tables, grouping by `category`.

6. Customer Orders by Ship Mode

- *Question:* How many orders were shipped by each shipping mode?
- **Hint***:* Use the `sales` table and group by `ship_mode`.

7. Sales by Date

- *Question:* What are the total sales for each month?
- **Hint***:* Use the `order_date` column from the `sales` table to group the data by month.

8. Customer Distribution by State

- *Question:* How many customers are there in each state?
- **Hint***:* Query the `customer` table and group the data by `state`.

9. Top 5 Customers by Sales

- *Question:* Who are the top 5 customers in terms of total sales?
- **Hint***:* Use the `sales` and `customer` tables to sum sales per customer, then sort by sales.

10. Product Performance in Subcategories

- *Question:* What is the total sales for each product subcategory?
- **Hint***:* Join the `sales` and `product` tables, grouping by `sub_category`.

11. Rank Products by Sales

- **Question:** How can we rank products by their total sales within each product category?
- **Hint:** Use the `RANK()` window function. You'll need to partition the data by `category` and order by total sales for each product.

```
RANK() OVER (PARTITION BY p.category ORDER BY SUM(s.sales) DESC)
```

12. Cumulative Sales by Date

- **Question:** How can we calculate cumulative sales over time (running total) for each product?
- **Hint:** Use the `SUM()` window function with an `ORDER BY` clause on `order_date` to create a running total for each product.

```
SUM(s.sales) OVER (PARTITION BY p.product_name ORDER BY s.order_date)
```

13. Find Top 3 Customers by Profit

- **Question:** How can we find the top 3 customers based on profit within each region?
- **Hint:** Use `RANK()` or `DENSE_RANK()` to assign ranks within each region based on total profit, and filter for the top 3 using `HAVING`.

```
RANK() OVER (PARTITION BY c.region ORDER BY SUM(s.profit) DESC)
```

14. Average Sales by Segment with Row Number

- **Question:** How can we find the average sales for each segment and assign a row number to each customer based on their sales?
- **Hint:** Use the `AVG()` window function to calculate average sales for each segment, and `ROW_NUMBER()` to assign a number to each row within the segment.

```
AVG(s.sales) OVER (PARTITION BY c.segment)
```

```
ROW_NUMBER() OVER (PARTITION BY c.segment ORDER BY s.sales DESC)
```

15. Difference in Sales Between Consecutive Days

- **Question:** How can we calculate the difference in sales between consecutive days for each product?
- **Hint:** Use `LAG()` to access the sales value from the previous day and subtract it from the current day's sales.

```
LAG(s.sales, 1, 0) OVER (PARTITION BY p.product_name ORDER BY
s.order_date)
```

16. Find Percent of Total Sales by Region

- **Question:** How can we calculate the percentage of total sales contributed by each region?
- **Hint:** Use the `SUM()` window function to calculate the total sales for all regions and then divide individual region sales by the total.

```
SUM(SUM(s.sales)) OVER ()
```

17. Calculate Moving Average of Sales

- **Question:** How can we calculate the moving average of sales over the last 3 orders for each product?
- **Hint:** Use `AVG()` with the `ROWS BETWEEN` clause to calculate the moving average over the previous 2 rows and the current row.

```
AVG(s.sales) OVER (PARTITION BY p.product_name ORDER BY s.order_date
ROWS BETWEEN 2 PRECEDING AND CURRENT ROW)
```

18. Find Largest and Smallest Order by Customer

- **Question:** How can we find the largest and smallest order (by sales) for each customer?
- **Hint:** Use `MAX()` and `MIN()` window functions to find the largest and smallest sales amounts for each customer.

```
MAX(s.sales) OVER (PARTITION BY c.customer_name)
```

```
MIN(s.sales) OVER (PARTITION BY c.customer_name)
```

19. Running Total of Profit by Customer

- **Question:** How can we calculate the running total of profit for each customer?
- **Hint:** Use `SUM()` with the `ORDER BY` clause to calculate the running total of profit for each customer based on their order date.

```
SUM(s.profit) OVER (PARTITION BY c.customer_name ORDER BY
s.order_date)
```

20. Calculate Dense Rank of Sales by Ship Mode

- **Question:** How can we assign a dense rank to each sale based on total sales, grouped by ship mode?
- **Hint:** Use the `DENSE_RANK()` function to assign ranks based on sales within each `ship_mode`.

```
DENSE_RANK() OVER (PARTITION BY s.ship_mode ORDER BY s.sales DESC)
```