

## Business Problem Statement

A major retail company aims to gain deeper insights into its customers' shopping behaviors to enhance sales performance, customer satisfaction, and long-term loyalty. Recently, the management has observed noticeable shifts in purchasing trends across different demographics, product categories, and sales channels (both online and offline). They seek to understand which key factors such as discounts, product reviews, seasonal influences, or preferred payment methods impact customer purchasing decisions and encourage repeat buying.

The objective of this analysis is to explore the company's consumer behavior data to uncover actionable trends that can help improve customer engagement, strengthen marketing efforts, and refine product and sales strategies.

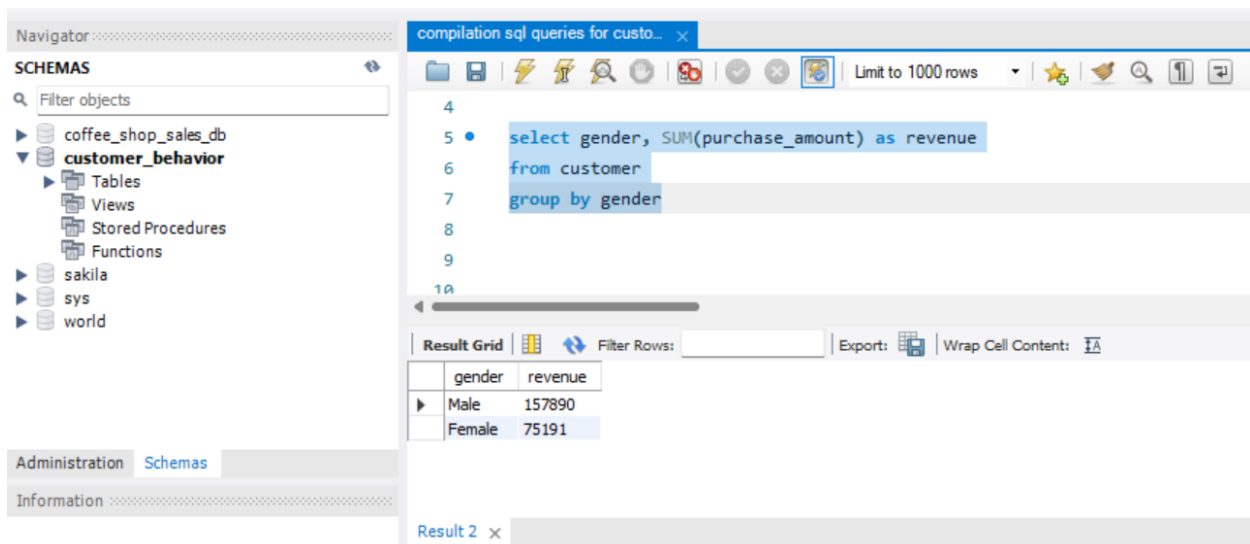
## SQL Queries

--Q1. What is the total revenue generated by male vs. female customers?

```
select gender, SUM(purchase_amount) as revenue
```

```
from customer
```

```
group by gender
```



The screenshot shows a SQL IDE interface. On the left, the 'SCHEMAS' pane displays a tree view with 'coffee\_shop\_sales\_db' expanded, showing 'customer\_behavior' (Tables, Views, Stored Procedures, Functions), 'sakila', 'sys', and 'world'. The main editor window, titled 'compilation sql queries for custo...', contains the following SQL query:

```
4
5 • select gender, SUM(purchase_amount) as revenue
6   from customer
7   group by gender
8
9
10
```

Below the query editor, the 'Result Grid' tab is active, displaying the results of the query in a table:

gender	revenue
Male	157890
Female	75191

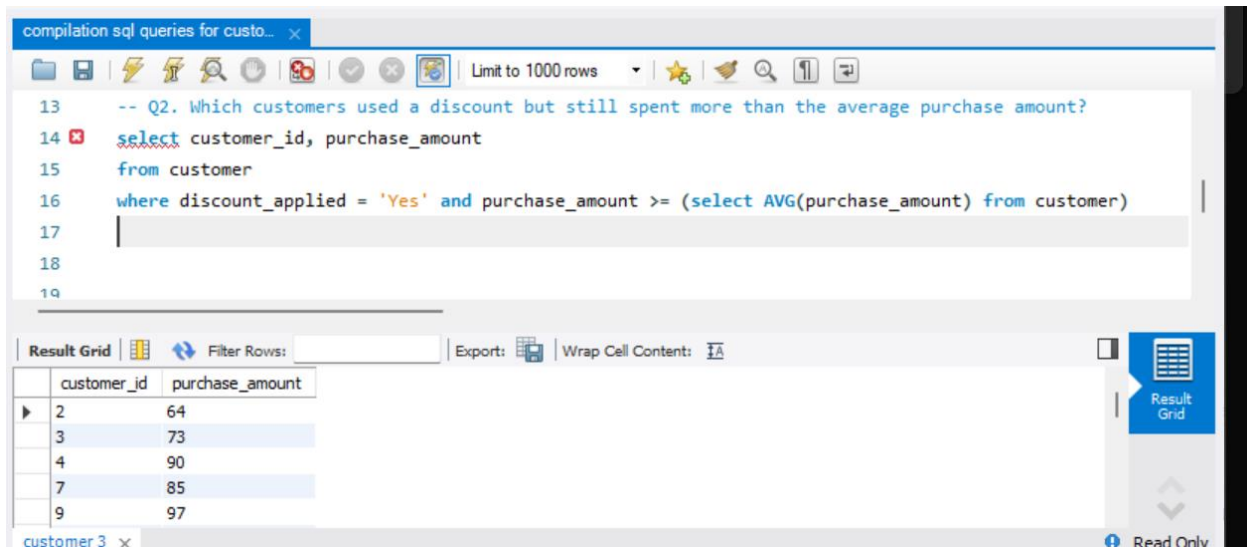
At the bottom, the 'Administration' tab is selected, showing 'Schemas' and 'Information' sections.

--Q2. Which customers used a discount but still spent more than the average purchase amount?

```
select customer_id, purchase_amount
```

```
from customer
```

```
where discount_applied = 'Yes' and purchase_amount >= (select AVG(purchase_amount) from customer)
```



The screenshot shows a SQL IDE window titled "compilation sql queries for custo...". The query editor contains the following SQL code:

```
13 -- Q2. Which customers used a discount but still spent more than the average purchase amount?
14 select customer_id, purchase_amount
15 from customer
16 where discount_applied = 'Yes' and purchase_amount >= (select AVG(purchase_amount) from customer)
17
18
19
```

Below the query editor, the "Result Grid" is displayed, showing the results of the query. The grid has two columns: "customer\_id" and "purchase\_amount". The results are as follows:

customer_id	purchase_amount
2	64
3	73
4	90
7	85
9	97

The IDE interface includes a toolbar with various icons, a "Limit to 1000 rows" dropdown, and a "Read Only" status indicator.

-- Q3. Which are the top 5 products with the highest average review rating?

```
select item_purchased , ROUND (AVG(review_rating),2) as "Average Product Rating"
```

```
from customer
```

```
group by item_purchased
```

```
order by avg(review_rating) desc
```

limit 5;

The screenshot shows a SQL IDE window titled "compilation sql queries for custo...". The query editor contains the following SQL code:

```
20 -- Q3. Which are the top 5 products with the highest average review rating?
21 select item_purchased , ROUND (AVG(review_rating),2) as "Average Product Rating"
22 from customer
23 group by item_purchased
24 order by avg(review_rating) desc
25 limit 5;
```

The "Result Grid" at the bottom displays the following data:

item_purchased	Average Product Rating
Gloves	3.86
Sandals	3.84
Boots	3.82
Hat	3.8

--Q4. Compare the average Purchase Amounts between Standard and Express Shipping.

```
select shipping_type,
ROUND(AVG(purchase_amount),2)
from customer
where shipping_type in ('Standard','Express')
group by shipping_type;
```

The screenshot shows a SQL IDE window with the following SQL code:

```
28 -- Q4. Compare the average Purchase Amounts between Standard and Express Shipping.
29 select shipping_type,
30 ROUND(AVG(purchase_amount),2)
31 from customer
32 where shipping_type in ('Standard','Express')
33 group by shipping_type;
```

The "Result Grid" at the bottom displays the following data:

shipping_type	ROUND(AVG(purchase_amount),2)
Express	60.48
Standard	58.46

--Q5. Do subscribed customers spend more? Compare average spend and total revenue

--between subscribers and non-subscribers.

```

SELECT subscription_status,
       COUNT(customer_id) AS total_customers,
       ROUND(AVG(purchase_amount),2) AS avg_spend,
       ROUND(SUM(purchase_amount),2) AS total_revenue
FROM customer
GROUP BY subscription_status
ORDER BY total_revenue,avg_spend DESC;

```

36 -- Q5. Do subscribed customers spend more? Compare average spend and total revenue between subscribers

37

38 • SELECT subscription\_status,

39       COUNT(customer\_id) AS total\_customers,

40       ROUND(AVG(purchase\_amount),2) AS avg\_spend,

41       ROUND(SUM(purchase\_amount),2) AS total\_revenue

42 FROM customer

43 GROUP BY subscription\_status

44 ORDER BY total\_revenue,avg\_spend DESC;

45

---

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

	subscription_status	total_customers	avg_spend	total_revenue
▶	Yes	1053	59.49	62645
	No	2847	59.87	170436

Result Grid

Form Editor

--Q6. Which 5 products have the highest percentage of purchases with discounts applied?

```

SELECT item_purchased,
       ROUND(100.0 * SUM(CASE WHEN discount_applied = 'Yes' THEN 1 ELSE 0 END)/COUNT(*),2)
AS discount_rate
FROM customer
GROUP BY item_purchased
ORDER BY discount_rate DESC
LIMIT 5;

```

```

47 -- Q6. Which 5 products have the highest percentage of purchases with discounts applied?
48
49 • SELECT item_purchased,
50        ROUND(100.0 * SUM(CASE WHEN discount_applied = 'Yes' THEN 1 ELSE 0 END)/COUNT(*),2) AS discount_
51 FROM customer
52 GROUP BY item_purchased
53 ORDER BY discount_rate DESC
54 LIMIT 5;
55
56
57

```

Result Grid

	item_purchased	discount_rate
▶	Hat	50.00
	Sneakers	49.66
	Coat	49.07
	Sweater	48.17
	Pants	47.37

Export: | Wrap Cell Content: | Fetch rows: | Result Grid | Form Editor

--Q7. Segment customers into New, Returning, and Loyal based on their total  
 -- number of previous purchases, and show the count of each segment.

```

with customer_type as (
SELECT customer_id, previous_purchases,
CASE
  WHEN previous_purchases = 1 THEN 'New'
  WHEN previous_purchases BETWEEN 2 AND 10 THEN 'Returning'
  ELSE 'Loyal'
END AS customer_segment
FROM customer)

select customer_segment, count(*) AS "Number of Customers"
from customer_type
group by customer_segment;

```

compilation sql queries for custo...

```

60 with customer_type as (
61   SELECT customer_id, previous_purchases,
62   CASE
63     WHEN previous_purchases = 1 THEN 'New'
64     WHEN previous_purchases BETWEEN 2 AND 10 THEN 'Returning'
65     ELSE 'Loyal'
66   END AS customer_segment
67 FROM customer)
68
69 select customer_segment, count(*) AS "Number of Customers"
70 from customer_type
71 group by customer_segment;

```

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

customer_segment	Number of Customers
Loyal	3116
Returning	701
New	83

Result Grid  
Form Editor

--Q8. What are the top 3 most purchased products within each category?

```

WITH item_counts AS (
  SELECT category,
         item_purchased,
         COUNT(customer_id) AS total_orders,
         ROW_NUMBER() OVER (PARTITION BY category ORDER BY COUNT(customer_id) DESC) AS
item_rank
  FROM customer
  GROUP BY category, item_purchased
)
SELECT item_rank, category, item_purchased, total_orders
FROM item_counts
WHERE item_rank <= 3;

```

```

75  -- Q8. What are the top 3 most purchased products within each category?
76
77  WITH item_counts AS (
78      SELECT category,
79             item_purchased,
80             COUNT(customer_id) AS total_orders,
81             ROW_NUMBER() OVER (PARTITION BY category ORDER BY COUNT(customer_id) DESC) AS item_rank
82      FROM customer
83      GROUP BY category, item_purchased
84  )
85  SELECT item_rank, category, item_purchased, total_orders
86  FROM item_counts
87  WHERE item_rank <= 3;

```

item_rank	category	item_purchased	total_orders
1	Accessories	Jewelry	171
2	Accessories	Sunglasses	161
3	Accessories	Belt	161
1	Clothing	Blouse	171
2	Clothing	Pants	171
3	Clothing	Shirt	169
1	Footwear	Sandals	160

--Q9. Are customers who are repeat buyers (more than 5 previous purchases) also likely to subscribe?

```

SELECT subscription_status,
       COUNT(customer_id) AS repeat_buyers
FROM customer
WHERE previous_purchases > 5
GROUP BY subscription_status;

```

```

91 -- Q9. Are customers who are repeat buyers (more than 5 previous purchases) also likely to subscribe?
92
93 • SELECT subscription_status,
94     COUNT(customer_id) AS repeat_buyers
95 FROM customer
96 WHERE previous_purchases > 5
97 GROUP BY subscription_status;
98

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
subscription_status	repeat_buyers			
Yes	958			
No	2518			

--Q10. What is the revenue contribution of each age group?

SELECT

age\_group,

SUM(purchase\_amount) AS total\_revenue

FROM customer

GROUP BY age\_group

ORDER BY total\_revenue desc;

```

100 -- Q10. What is the revenue contribution of each age group?
101
102 • SELECT
103     age_group,
104     SUM(purchase_amount) AS total_revenue
105 FROM customer
106 GROUP BY age_group
107 ORDER BY total_revenue desc;
108

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
age_group	total_revenue			
Young Adult	62143			
Middle-aged	59197			
Adult	55978			
Senior	55763			