

# **NEW WHEELS PROJECT**

## **1. PROBLEM STATEMENT:**

A lot of people in the world share a common desire: to own a vehicle. A car or an automobile is seen as an object that gives the freedom of mobility. Many now prefer pre-owned vehicles because they come at an affordable cost, but at the same time, they are also concerned about whether the after-sales service provided by the resale vendors is as good as the care you may get from the actual manufacturers. New-Wheels, a vehicle resale company, has launched an app with an end-to-end service from listing the vehicle on the platform to shipping it to the customer's location. This app also captures the overall after-sales feedback given by the customer.

## **2. OBJECTIVE:**

New-Wheels sales have been dipping steadily in the past year, and due to the critical customer feedback and ratings online, there has been a drop in new customers every quarter, which is concerning to the business. The CEO of the company now wants a quarterly report with all the key metrics sent to him so he can assess the health of the business and make the necessary decisions.

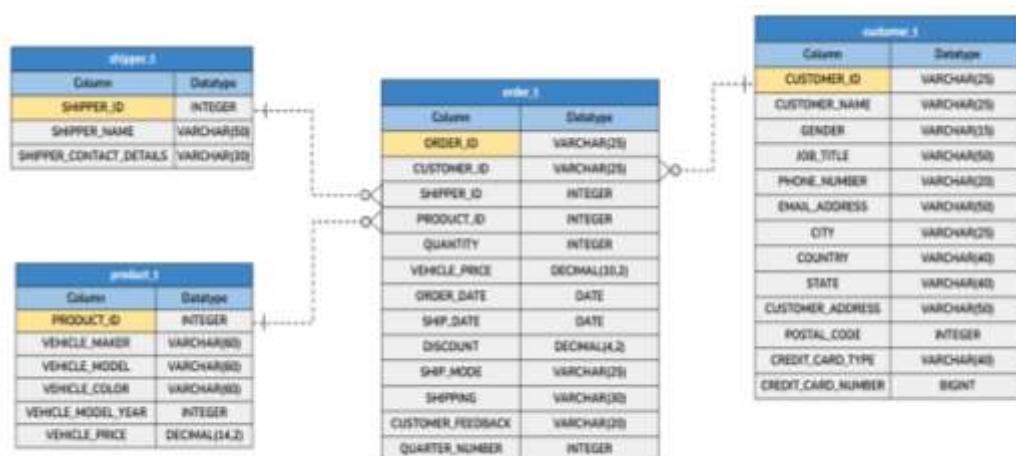
As a data analyst, you see that there is an array of questions that are being asked at the leadership level that need to be answered using data. Import the dump file that contains various tables that are present in the database. Use the data to answer the questions posed and create a quarterly business report for the CEO.

## **3. DATA DICTIONARY:**

- shipper\_id: Unique ID of the Shipper
- shipper\_name: Name of the Shipper
- shipper\_contact\_details: Contact detail of the Shipper
- product\_id: Unique ID of the Product
- vehicle\_maker: Vehicle Manufacturing company name
- vehicle\_model: Vehicle model name
- vehicle\_color: Color of the Vehicle
- vehicle\_model\_year: Year of Manufacturing
- vehicle\_price: Price of the Vehicle
- quantity: Ordered Quantity

- customer\_id: Unique ID of the customer
- customer\_name: Name of the customer
- gender: Gender of the customer
- job\_title: Job Title of the customer
- phone\_number: Contact detail of the customer
- email\_address: Email address of the customer
- city: Residing city of the customer
- country: Residing country of the customer
- state: Residing state of the customer
- customer\_address: Address of the customer
- order\_date: Date on which customer ordered the vehicle
- order\_id: Unique ID of the order
- ship\_date: Shipment Date
- ship\_mode: Shipping Mode/Class
- shipping: Shipping Ways
- postal\_code: Postal Code of the customer
- discount: Discount given to the customer for the particular order by credit card in percentage4.
- credit\_card\_type: Credit Card Type
- credit\_card\_number: Credit card number
- customer\_feedback: Feedback of the customer
- quarter\_number : Quarter Number

## 4. ER DIAGRAM:



## **BUSINESS QUESTIONS:**

1(a). Find the total number of customers who have placed orders.

### **Solution Query & Output:**

```
SELECT COUNT(DISTINCT Customer_ID) AS TOTAL_ACTIVE_CUSTOMERS  
FROM Order_t;
```



The screenshot shows a database query result in a web interface. The top section displays the query result in a table with one column, 'TOTAL\_ACTIVE\_CUSTOMERS', and one row with the value '994'. Below the table, there is a 'Result' tab and an 'Output' tab. The 'Output' tab is active, showing a log of the query execution. The log entry indicates that the query 'SELECT COUNT(DISTINCT CUSTOMER\_ID) AS TOTAL\_ACTIVE\_CUSTOMERS' was executed at 21:35:17, returned 1 row(s), and took 0.047 seconds to execute.

Result
TOTAL_ACTIVE_CUSTOMERS
994

#	Time	Action	Message	Duration / Fetch
4	21:35:17	SELECT COUNT(DISTINCT CUSTOMER_ID) AS TOTAL_ACTIVE_CUSTOMERS	1 row(s) returned	0.047 sec / 0.000 sec

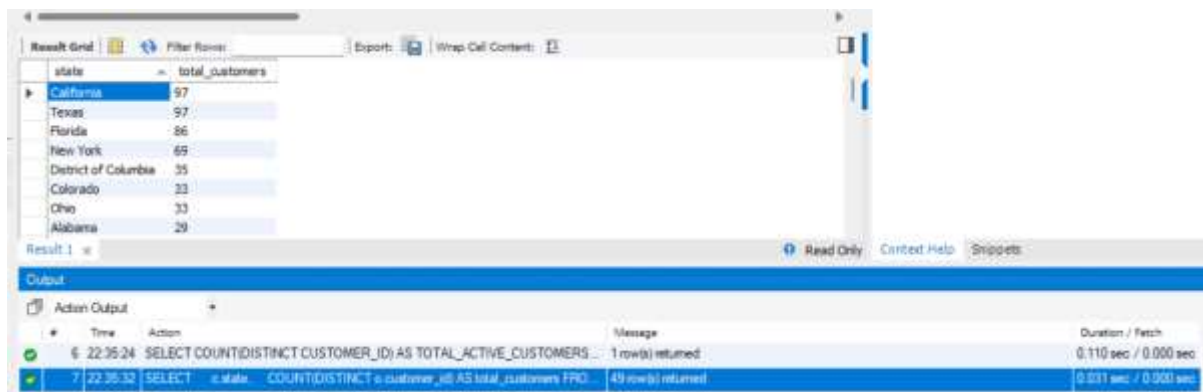
### **Observation & Insights:**

- A total of 994 unique customers have placed orders, indicating a strong active customer base.

1(b). What is the distribution of the customers across states?

### Solution Query & Output:

```
SELECT
  c.state,
  COUNT(DISTINCT o.customer_id) AS total_customers
FROM order_t o
JOIN customer_t c
  ON o.customer_id = c.customer_id
GROUP BY c.state
ORDER BY total_customers DESC;
```



The screenshot shows a database query result grid with two columns: 'state' and 'total\_customers'. The results are ordered by total\_customers in descending order. The states and their corresponding customer counts are: California (97), Texas (97), Florida (86), New York (69), District of Columbia (35), Colorado (33), Ohio (33), and Alabama (29). Below the grid, the 'Output' section shows the execution details of the query, including the time taken and the number of rows returned.

state	total_customers
California	97
Texas	97
Florida	86
New York	69
District of Columbia	35
Colorado	33
Ohio	33
Alabama	29

#	Time	Action	Message	Duration / Fetch
6	22:35:24	SELECT COUNT(DISTINCT CUSTOMER_ID) AS TOTAL_ACTIVE_CUSTOMERS	1 row(s) returned	0.110 sec / 0.000 sec
7	22:35:32	SELECT c.state, COUNT(DISTINCT o.customer_id) AS total_customers FROM	49 row(s) returned	0.031 sec / 0.000 sec

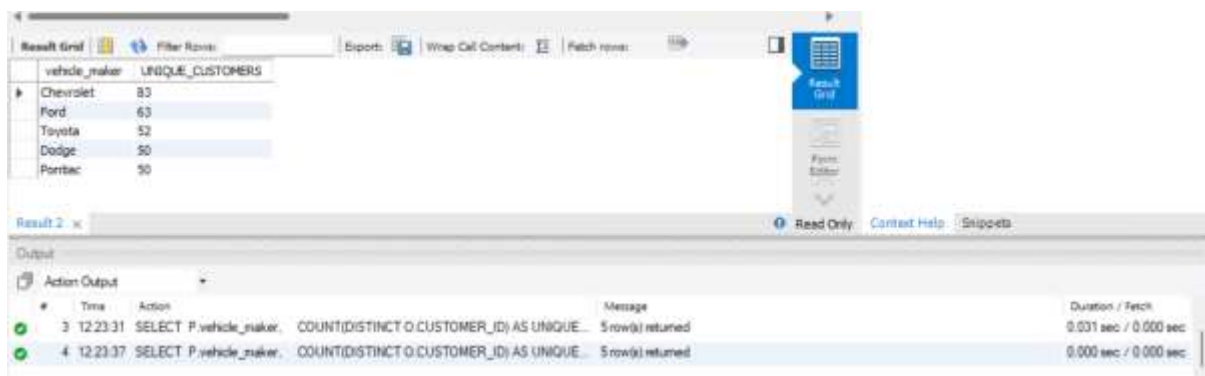
### Observation & Insights:

- The highest number of customers are concentrated in **California and Texas (97 each)**, followed by **Florida (86)** and **New York (69)**, indicating strong customer engagement and higher market penetration in these key states.
- States such as **Maine, Vermont, and Wyoming** show minimal customer presence, highlighting limited market reach.

2. Which are the top 5 vehicle makers preferred by the customers?

### Solution Query & Output:

```
SELECT
    p.vehicle_maker,
    COUNT(DISTINCT O.CUSTOMER_ID) AS UNIQUE_CUSTOMERS
FROM order_t o
JOIN product_t p
    ON o.product_id = p.product_id
GROUP BY p.vehicle_maker
ORDER BY unique_customers DESC
LIMIT 5;
```



vehicle_maker	UNIQUE_CUSTOMERS
Chevrolet	83
Ford	63
Toyota	52
Dodge	50
Pontiac	50

#	Time	Action	Message	Duration / Fetch
3	12:23:31	SELECT P.vehicle_maker, COUNT(DISTINCT O.CUSTOMER_ID) AS UNIQUE...	5 row(s) returned	0.031 sec / 0.000 sec
4	12:23:37	SELECT P.vehicle_maker, COUNT(DISTINCT O.CUSTOMER_ID) AS UNIQUE...	5 row(s) returned	0.000 sec / 0.000 sec

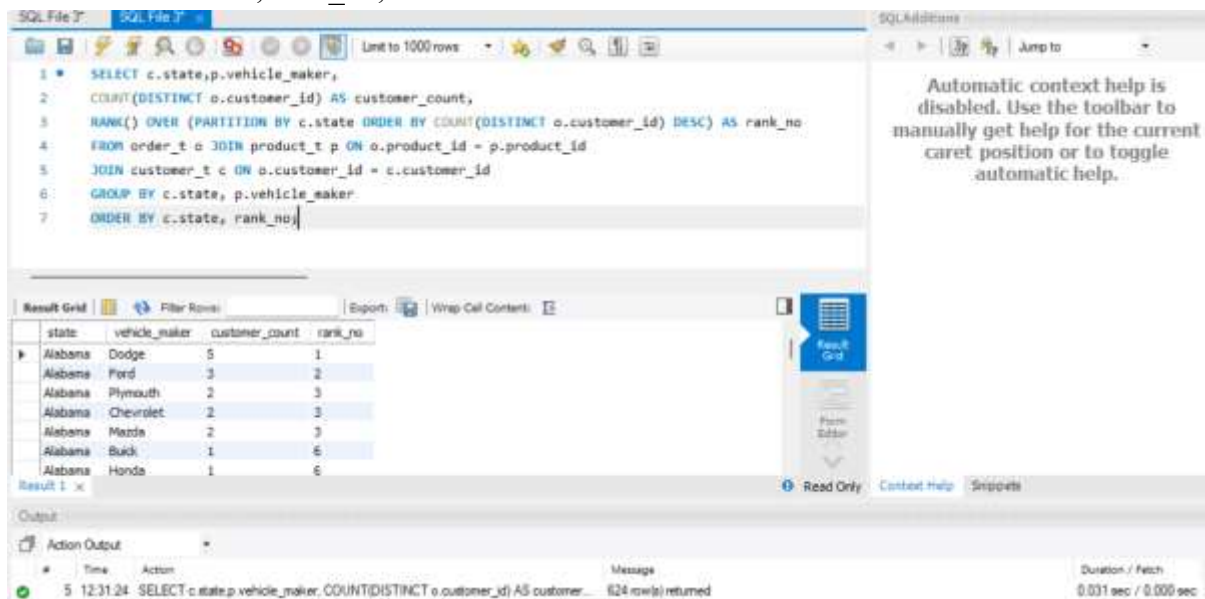
### Observation & Insights:

- **Chevrolet** is the most preferred vehicle maker with **83 unique customers**, indicating strong customer demand and brand popularity.
- **Ford** ranks second with **63 customers**, followed by **Toyota (52)**, **Pontiac (50)**, and **Dodge (50)**, reflecting steady demand across multiple brands.
- Customer preference is evenly spread across other makers, showing no strong brand dominance.

3. Which is the most preferred vehicle maker in each state?

### Solution Query & Output:

```
SELECT c.state,p.vehicle_maker,  
COUNT(DISTINCT o.customer_id) AS customer_count,  
RANK() OVER (PARTITION BY c.state ORDER BY COUNT(DISTINCT o.customer_id)  
DESC) AS rank_no  
FROM order_t o JOIN product_t p ON o.product_id = p.product_id  
JOIN customer_t c ON o.customer_id = c.customer_id  
GROUP BY c.state, p.vehicle_maker  
ORDER BY c.state, rank_no;
```



The screenshot shows the SQL Developer interface. The top pane displays the SQL query. The bottom pane shows the 'Result Grid' with the following data:

state	vehicle_maker	customer_count	rank_no
Alabama	Dodge	5	1
Alabama	Ford	3	2
Alabama	Plymouth	2	3
Alabama	Chevrolet	2	3
Alabama	Mazda	2	3
Alabama	Buick	1	6
Alabama	Honda	1	6

The bottom pane also shows the 'Output' section with the following message:

```
5 12:31:24 SELECT c.state,p.vehicle_maker,COUNT(DISTINCT o.customer_id) AS customer_... 624 row(s) returned  
Duration / Fetch: 0.031 sec / 0.000 sec
```

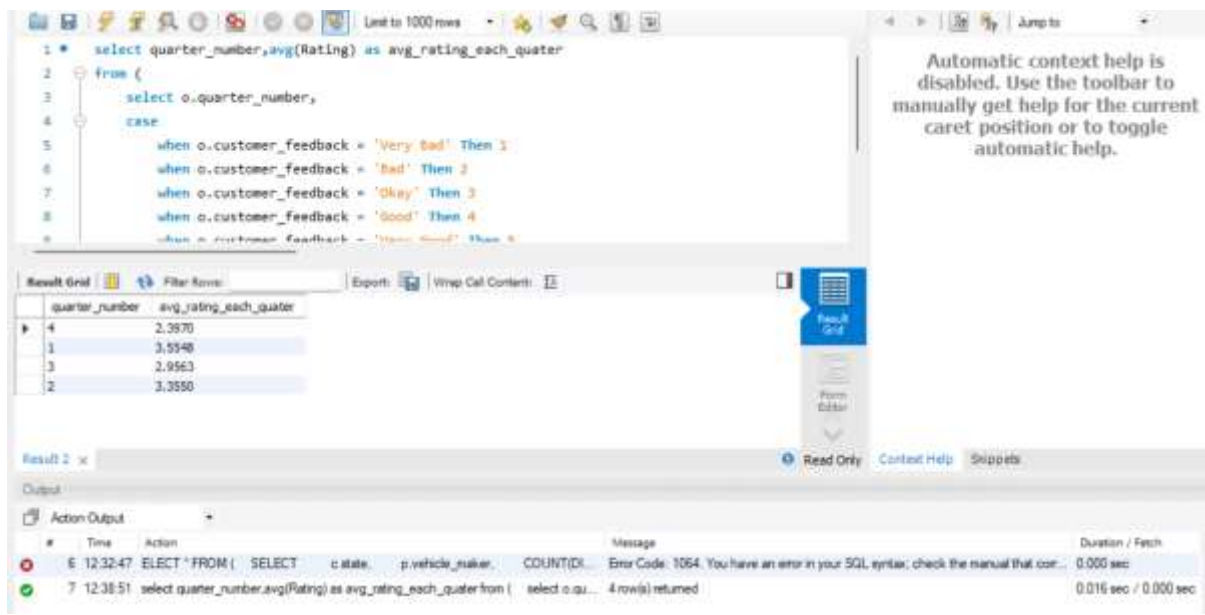
### Observation & Insights:

- Chevrolet emerges as the most preferred brand across several states, including Texas, Colorado, Ohio, and Washington, indicating strong nationwide demand and brand presence.
- Toyota, Ford, Dodge, and Pontiac also rank among the top choices in multiple states, reflecting broad customer trust and consistent market acceptance across regions.
- In smaller states with fewer orders, customer preferences appear more evenly distributed, with multiple brands sharing the top position due to lower purchase volumes.

4. Find the overall average rating given by the customers. What is the average rating in each quarter?

### Solution Query & Output:

```
select quarter_number, avg(Rating) as avg_rating_each_quater
from (
    select o.quarter_number,
    case
        when o.customer_feedback = 'Very Bad' Then 1
        when o.customer_feedback = 'Bad' Then 2
        when o.customer_feedback = 'Okay' Then 3
        when o.customer_feedback = 'Good' Then 4
        when o.customer_feedback = 'Very Good' Then 5
    else null
    end as Rating
    from order_t o
) as feedback_ratings
group by quarter_number;
```



The screenshot shows a SQL IDE interface. The top pane displays a SQL query that calculates the average rating for each quarter based on customer feedback. The bottom pane shows the results of the query in a table format. The table has two columns: 'quarter\_number' and 'avg\_rating\_each\_quater'. The results are as follows:

quarter_number	avg_rating_each_quater
4	2.3970
1	3.5548
3	2.9563
2	3.3550

Below the table, there is a 'Result 2' tab and an 'Output' section. The 'Output' section shows a message: 'Error Code: 1064. You have an error in your SQL syntax; check the manual that corresponds to your MySQL server version for the right syntax to use near 'select o.qu...' at line 1'. The message also indicates that 4 row(s) were returned.

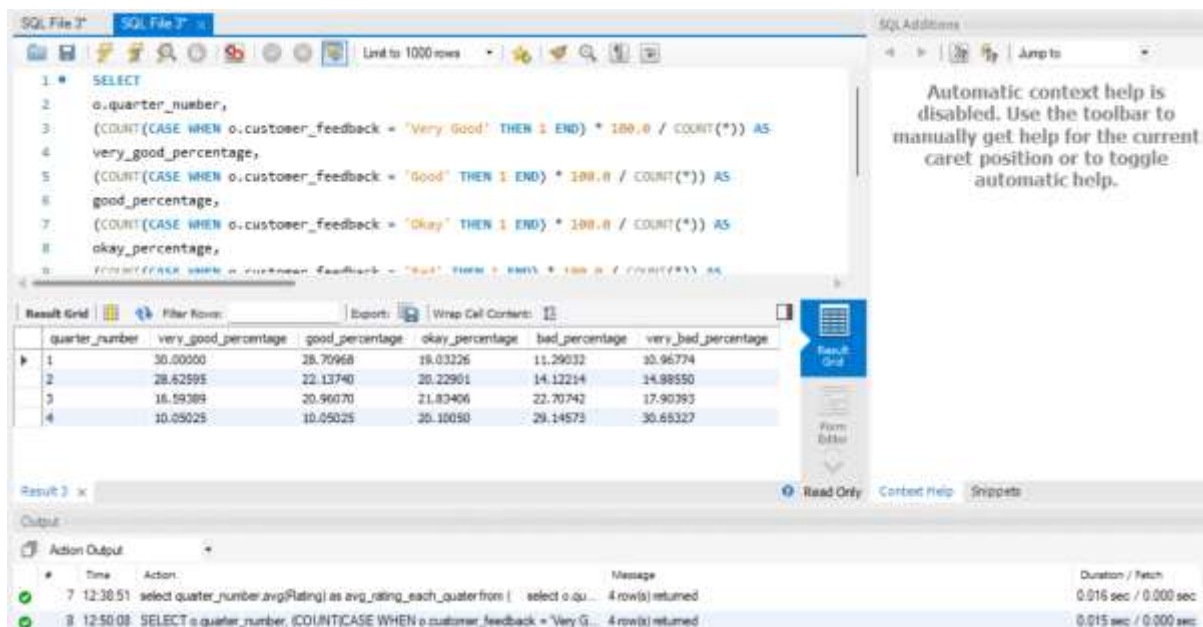
### Observation & Insights:

- Overall customer satisfaction remains **moderate**, with quarterly average ratings fluctuating between **2.4 and 3.6**, indicating inconsistent service quality throughout the year.
- **Quarter 1 recorded the highest average rating (3.55)**, reflecting a positive customer experience and stronger service performance during the early part of the year.
- **Quarter 4 showed the lowest rating (2.39)**, suggesting a noticeable decline in customer satisfaction toward year-end.
- Addressing Q4 issues and applying Q1 best practices can improve overall satisfaction.

5. Find the percentage distribution of feedback from the customers. Are customers getting more dissatisfied over time?

### Solution Query & Output:

```
SELECT
o.quarter_number,
(COUNT(CASE WHEN o.customer_feedback = 'Very Good' THEN 1 END) * 100.0 /
COUNT(*)) AS
very_good_percentage,
(COUNT(CASE WHEN o.customer_feedback = 'Good' THEN 1 END) * 100.0 / COUNT(*))
AS
good_percentage,
(COUNT(CASE WHEN o.customer_feedback = 'Okay' THEN 1 END) * 100.0 / COUNT(*))
AS
okay_percentage,
(COUNT(CASE WHEN o.customer_feedback = 'Bad' THEN 1 END) * 100.0 / COUNT(*))
AS
bad_percentage,
(COUNT(CASE WHEN o.customer_feedback = 'Very Bad' THEN 1 END) * 100.0 /
COUNT(*)) AS
very_bad_percentage
FROM order_t o
GROUP BY o.quarter_number
ORDER BY o.quarter_number;
```



The screenshot shows a SQL IDE window with a query editor and a results grid. The query is the same as the one provided in the text. The results grid displays the output for 4 quarters, showing the percentage distribution of customer feedback for each quarter.

quarter_number	very_good_percentage	good_percentage	okay_percentage	bad_percentage	very_bad_percentage
1	30.00000	28.70968	19.03226	11.29032	20.96774
2	28.62595	22.13740	20.22901	14.12214	14.98550
3	16.59389	20.96070	21.83406	22.70742	17.90793
4	10.05025	10.05025	20.10050	29.14573	30.65327

The bottom of the screenshot shows the 'Output' pane with a log of SQL actions and their execution times.

#	Time	Action	Message	Duration / Fetch
7	12:30:51	select quarter_number,avg(Rating) as avg_rating_each_quater from	select o.qu... 4 row(s) returned	0.016 sec / 0.000 sec
8	12:50:08	SELECT o.quarter_number, (COUNT(CASE WHEN o.customer_feedback = 'Very G...	4 row(s) returned	0.015 sec / 0.000 sec



### **Observation & Insights:**

- Customer satisfaction shows a **consistent decline across quarters**, indicating a gradual deterioration in overall service quality over time.
- In **Quarter 1**, nearly **59% of customers provided “Good” or “Very Good” feedback**, reflecting strong customer satisfaction and a positive service experience.
- By **Quarter 4**, positive feedback fell sharply to around **20%**, while negative feedback increased to nearly **60%**, highlighting a significant drop in customer perception.
- This noticeable shift toward negative sentiment suggests growing dissatisfaction, which may contribute to lower customer retention and declining sales.
- It is recommended that the company **closely review operations during Q3 and Q4**, identify key service gaps, and implement corrective measures to restore customer trust and improve overall experience.

6. What is the trend of the number of orders by quarter?

### Solution Query & Output:

```
SELECT o.quarter_number,COUNT(o.order_id) AS total_orders
FROM order_t o
GROUP BY o.quarter_number
ORDER BY o.quarter_number;
```

The screenshot displays a database query interface. At the top, the SQL query is entered in a text area:

```
1 SELECT o.quarter_number,COUNT(o.order_id) AS total_orders
2 FROM order_t o
3 GROUP BY o.quarter_number
4 ORDER BY o.quarter_number;
```

Below the query, a "Result Grid" shows the output of the query. The grid has two columns: "quarter\_number" and "total\_orders". The data is as follows:

quarter_number	total_orders
1	310
2	262
3	229
4	199

At the bottom, an "Action Output" table shows the execution details of the query:

Time	Action	Message	Duration / Fetch
12:50:08	SELECT o.quarter_number, COUNT(CASE WHEN o.customer_feedback = 'Very G...	4 row(s) returned	0.015 sec / 0.000 sec
13:00:45	SELECT o.quarter_number,COUNT(o.order_id) AS total_orders FROM order_t o GR...	4 row(s) returned	0.000 sec / 0.000 sec

### Observation & Insights:

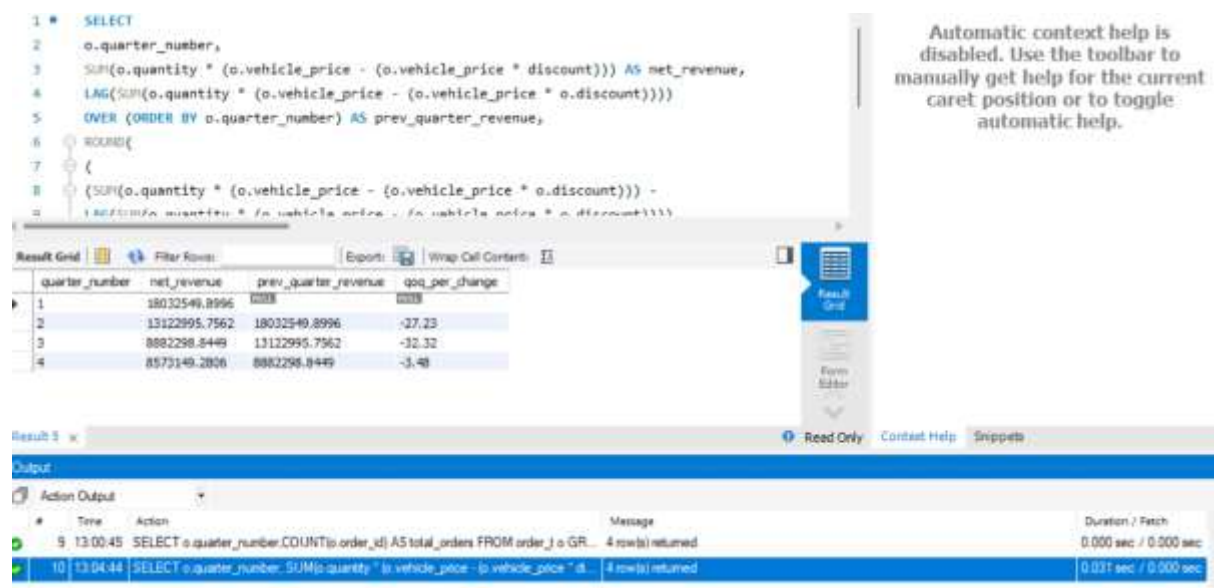
- **Quarter 1** recorded the highest orders (310), indicating strong demand at the start of the year.
- Orders declined steadily through **Q2 (262)** and **Q3 (229)**, showing weakening customer engagement.
- **Quarter 4** reported the lowest orders (199), marking the poorest performance of the year.
- The continuous downward trend suggests possible issues with customer satisfaction and repeat purchases, requiring immediate management action.

7. Calculate the net revenue generated by the company. What is the quarter-over-quarter % change in net revenue?

### Solution Query & Output:

```
SELECT
o.quarter_number,
SUM(o.quantity * (o.vehicle_price - (o.vehicle_price * discount))) AS net_revenue,
LAG(SUM(o.quantity * (o.vehicle_price - (o.vehicle_price * o.discount))))
OVER (ORDER BY o.quarter_number) AS prev_quarter_revenue,
ROUND(
(
(SUM(o.quantity * (o.vehicle_price - (o.vehicle_price * o.discount))) -
LAG(SUM(o.quantity * (o.vehicle_price - (o.vehicle_price * o.discount))))
OVER (ORDER BY o.quarter_number))
/ LAG(SUM(o.quantity * (o.vehicle_price - (o.vehicle_price * o.discount))))
OVER (ORDER BY o.quarter_number)
) * 100, 2
) AS qoq_per_change
FROM order_t o
GROUP BY o.quarter_number
ORDER BY o.quarter_number;
```

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.



The screenshot displays a SQL query editor with the following query:

```
1 * SELECT
2   o.quarter_number,
3   SUM(o.quantity * (o.vehicle_price - (o.vehicle_price * discount))) AS net_revenue,
4   LAG(SUM(o.quantity * (o.vehicle_price - (o.vehicle_price * o.discount))))
5   OVER (ORDER BY o.quarter_number) AS prev_quarter_revenue,
6   ROUND(
7   (
8   (SUM(o.quantity * (o.vehicle_price - (o.vehicle_price * o.discount))) -
9   LAG(SUM(o.quantity * (o.vehicle_price - (o.vehicle_price * o.discount))))
10  OVER (ORDER BY o.quarter_number))
11  / LAG(SUM(o.quantity * (o.vehicle_price - (o.vehicle_price * o.discount))))
12  OVER (ORDER BY o.quarter_number)
13  ) * 100, 2
14  ) AS qoq_per_change
15  FROM order_t o
16  GROUP BY o.quarter_number
17  ORDER BY o.quarter_number;
```

The results are shown in a table with the following data:

quarter_number	net_revenue	prev_quarter_revenue	qoq_per_change
1	18032549.8996	NULL	NULL
2	13122995.7562	18032549.8996	-27.23
3	8882298.5449	13122995.7562	-32.32
4	8573149.2806	8882298.5449	-3.48

The bottom section of the screenshot shows the 'Action Output' pane with the following messages:

#	Time	Action	Message	Duration / Patch
9	13:00:45	SELECT o.quarter_number,COUNT(o.order_id) AS total_orders FROM order_t o GR...	4 row(s) returned	0.000 sec / 0.000 sec
10	13:04:44	SELECT o.quarter_number,SUM(o.quantity * (o.vehicle_price - (o.vehicle_price * d...	4 row(s) returned	0.031 sec / 0.000 sec

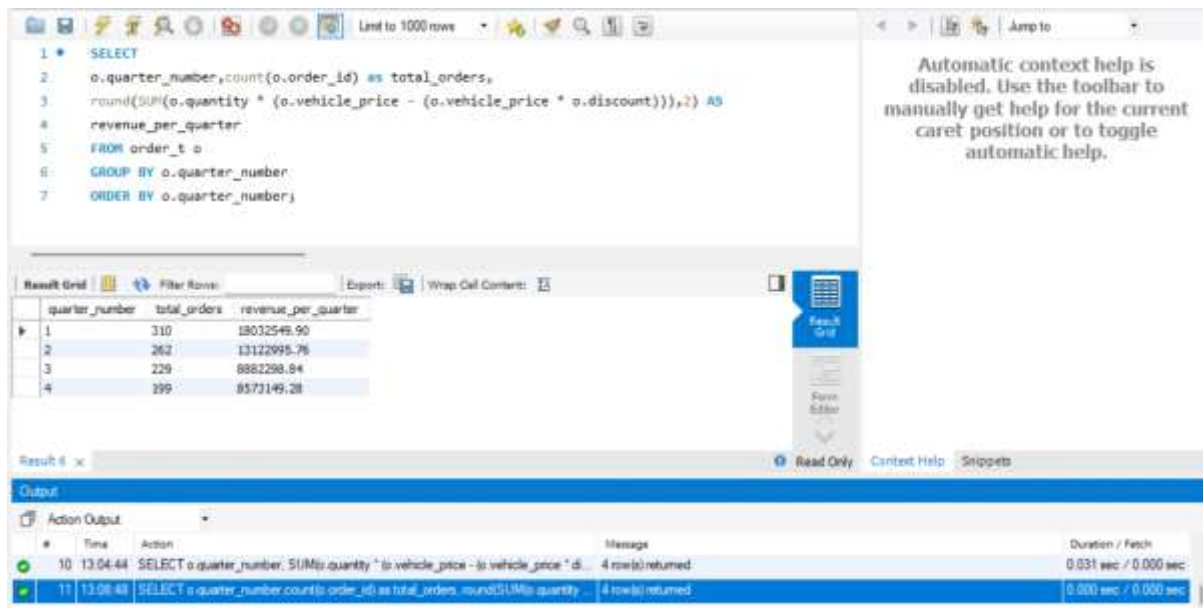
### **Observation & Insights:**

- **Quarter 1** generated the highest revenue, indicating a strong start to the year.
- **Quarter 2** experienced a ~27% decline, signaling the beginning of a downward trend.
- **Quarter 3** recorded the steepest drop (~32%), reflecting reduced sales and weaker customer engagement.
- **Quarter 4** showed only a marginal decline (~3%), suggesting the slowdown stabilized but did not recover.
- Overall, the consistent fall in revenue aligns with declining order volumes and customer satisfaction observed earlier.

8. What is the trend of net revenue and orders by quarters?

### Solution Query & Output:

```
SELECT
o.quarter_number,count(o.order_id) as total_orders,
round(SUM(o.quantity * (o.vehicle_price - (o.vehicle_price * o.discount))),2) AS
revenue_per_quarter
FROM order_t o
GROUP BY o.quarter_number
ORDER BY o.quarter_number;
```



Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

quarter_number	total_orders	revenue_per_quarter
1	310	18032549.90
2	262	13122995.76
3	229	8882298.84
4	199	8573149.28

Result 6 x

Read Only Context Help Snippets

Output

#	Time	Action	Message	Duration / Fetch
10	13:04:44	SELECT o.quarter_number, SUM(o.quantity * (o.vehicle_price - (o.vehicle_price * o.discount))) AS	4 row(s) returned	0.031 sec / 0.000 sec
11	13:04:48	SELECT o.quarter_number, count(o.order_id) as total_orders, round(SUM(o.quantity * (o.vehicle_price - (o.vehicle_price * o.discount))),2) AS	4 row(s) returned	0.000 sec / 0.000 sec

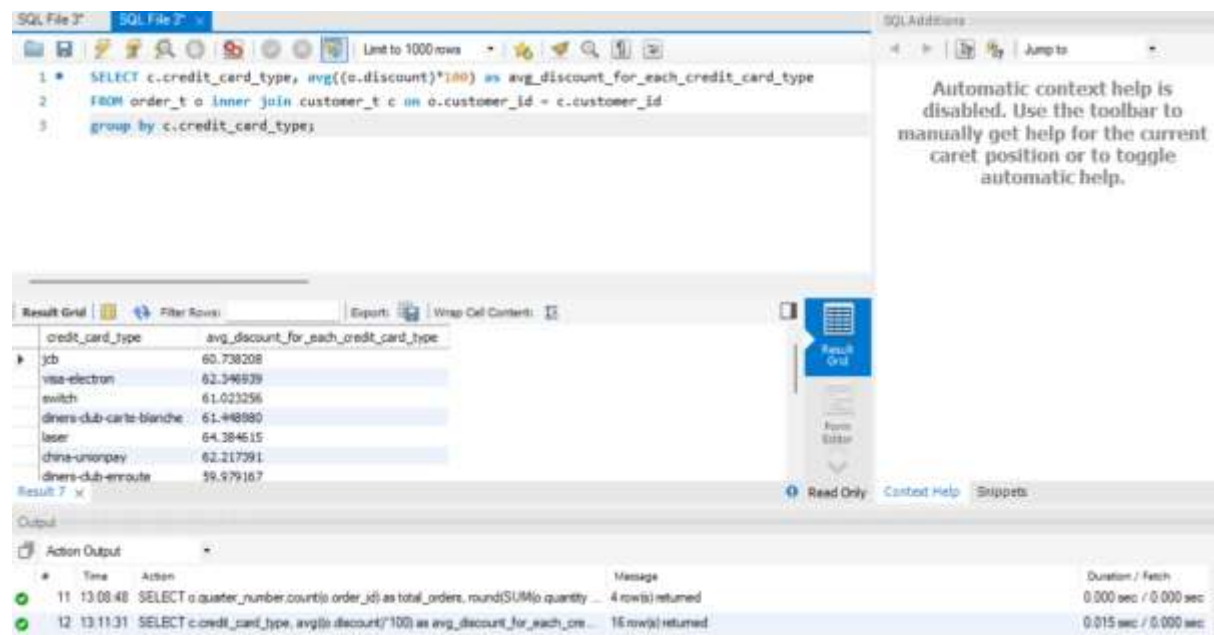
### Observation & Insights:

- Both **order volume** and **net revenue** declined consistently across quarters, indicating a steady slowdown in business performance.
- **Quarter 1** delivered the strongest results, with **310 orders** generating **₹18.03M** in revenue.
- **Quarter 4** recorded the weakest performance, with only **199 orders** and **₹8.57M** in revenue.
- The sustained decline suggests reduced customer demand and engagement over time.

9. What is the average discount offered for different types of credit cards?

### Solution Query & Output:

```
SELECT c.credit_card_type, avg((o.discount)*100) as  
avg_discount_for_each_credit_card_type  
FROM order_t o inner join customer_t c on o.customer_id = c.customer_id  
group by c.credit_card_type;
```



The screenshot shows the SQL Developer interface. The top pane displays the SQL query. The bottom pane shows the results in a grid format. The results table has two columns: 'credit\_card\_type' and 'avg\_discount\_for\_each\_credit\_card\_type'. The data rows are as follows:

credit_card_type	avg_discount_for_each_credit_card_type
jcb	60.738208
visa-electron	62.346939
switch	61.023256
diners-club-carte-blanche	61.448890
laser	64.384615
china-unionpay	62.217391
diners-club-enroute	59.979167

The bottom pane also shows the execution log with two entries:

#	Time	Action	Message	Duration / Fetch
11	13:08:48	SELECT o.quarter_number count(o.order_id) as total_orders, round(SUM(o.quantity ...	4 row(s) returned	0.000 sec / 0.000 sec
12	13:11:31	SELECT c.credit_card_type, avg(o.discount*100) as avg_discount_for_each_cre ...	16 row(s) returned	0.015 sec / 0.000 sec

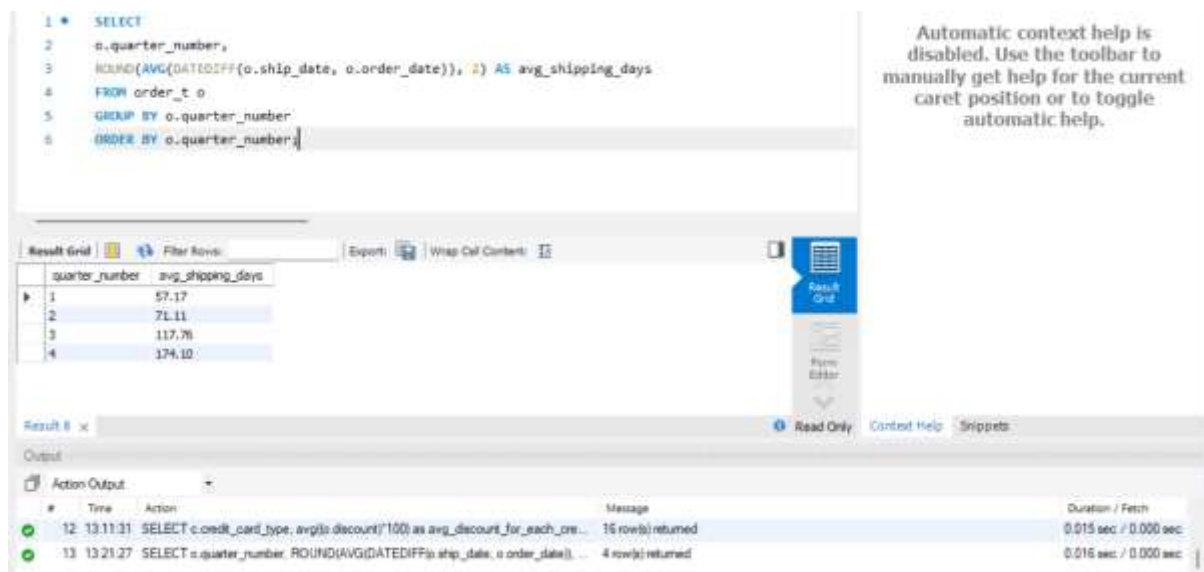
### Observation & Insights:

- The average discount remains relatively stable across all credit card types, ranging between **58% and 63%**.
- **Visa Electron, Mastercard, and China UnionPay** customers receive slightly higher average discounts compared to other cardholders.
- **Diners Club International and Solo** show the lowest average discount levels, at around **58%**.
- Overall, discount variations are minimal, indicating a largely consistent pricing strategy across payment methods.

10. What is the average time taken to ship the placed orders for each quarter?

**Solution Query & Output:**

```
SELECT
o.quarter_number,
ROUND(AVG(DATEDIFF(o.ship_date, o.order_date)), 2) AS avg_shipping_days
FROM order_t o
GROUP BY o.quarter_number
ORDER BY o.quarter_number;
```



Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

quarter_number	avg_shipping_days
1	57.17
2	71.11
3	117.76
4	174.10

Result 8: x

Output

#	Time	Action	Message	Duration / Fetch
12	13:11:31	SELECT c.credit_card_type, avg(is_discount/100) as avg_discount_for_each_pre...	16 row(s) returned	0.015 sec / 0.000 sec
13	13:21:27	SELECT o.quarter_number, ROUND(AVG(DATEDIFF(o.ship_date, o.order_date)), ...	4 row(s) returned	0.016 sec / 0.000 sec

**Observation & Insights:**

- The **average shipping time increased consistently across quarters**, indicating progressively slower deliveries over time.
- **Quarter 1** recorded the fastest delivery performance at approximately **57 days**, whereas **Quarter 4** experienced the longest delays at around **174 days**.
- This upward trend suggests inefficiencies in logistics or order processing.
- The company should review and optimize its shipping operations to enhance delivery speed and overall customer experience

## **BUSINESS METRICS:**

Total Revenue	Total Orders	Total Customers	Average Rating
48610993.78	1000	994	3
Last Quarter Revenue	Last quarter Orders	Average Days to Ship	% Good Feedback
8573149.28	199	97.96	21.5

## **BUSINESS RECOMMENDATIONS:**

- Customer satisfaction requires immediate attention, with an average rating of only **3.0** and just **21.5% of customers providing positive (“Good”) feedback**.
- **Shipping delays remain high (≈98 days)**, negatively impacting customer experience; optimizing logistics and delivery processes is critical.
- The sharp decline in **orders and revenue in the last quarter** signals weakening demand and reduced customer engagement.
- Analyzing customer feedback can help identify key pain points and enable targeted service improvements.
- Introducing **loyalty programs, personalized offers, and discounts** can improve retention and encourage repeat purchases.
- Overall, enhancing delivery efficiency, customer experience, and promotional strategies will be essential to drive future growth and restore sales momentum.