New Wheels Project

Introduction to SQL

Problem Statement

Business Context

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.

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.

Business Questions

Question 1: Find the total number of customers who have placed orders. What is the distribution of the customers across states?

Solution Query:

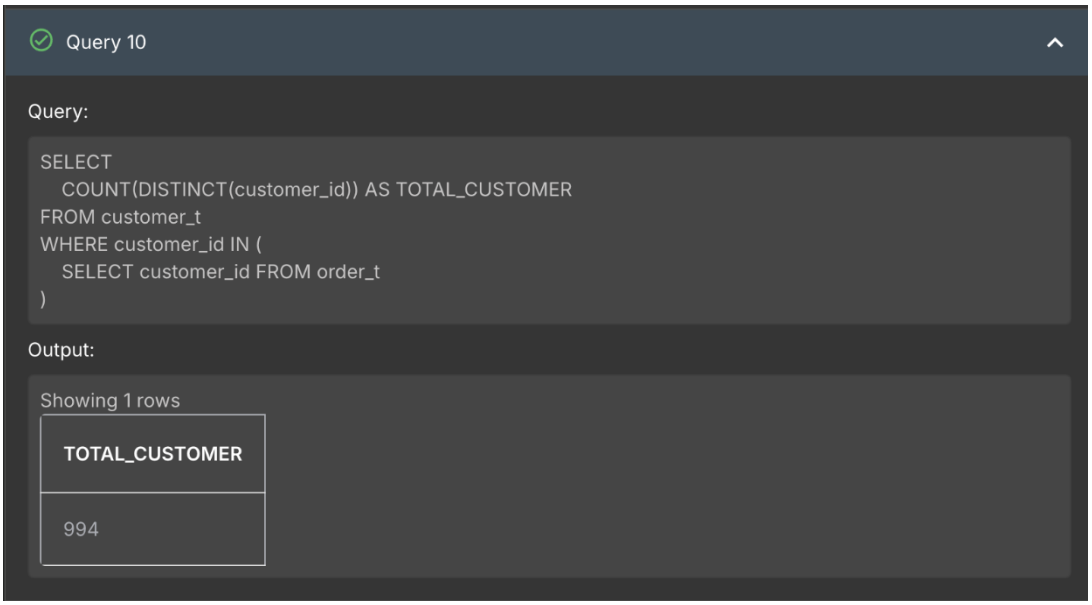
Find the total number of customers who have placed orders.

```
SELECT
    COUNT(DISTINCT(customer_id)) AS TOTAL_CUSTOMER
FROM customer_t
WHERE customer_id IN (
    SELECT customer_id FROM order_t
);
```

What is the distribution of the customers across states?

```
SELECT
    state,
    COUNT(customer_id) AS NUMBER_OF_CUSTOMER
FROM customer_t
GROUP BY state
ORDER BY NUMBER_OF_CUSTOMER DESC;
```

Output:



Result: Passed

Query 1

Query:

```
SELECT
   state,
   COUNT(customer_id) AS NUMBER_OF_CUSTOMER
FROM customer_t
GROUP BY state
ORDER BY NUMBER_OF_CUSTOMER DESC
```

Output:

Showing first 10 rows out of 49 rows

state	NUMBER_OF_CUSTOM...
Texas	97
California	97
Florida	86
New York	69
District of Columbia	35
Ohio	33

Observations and Insights:

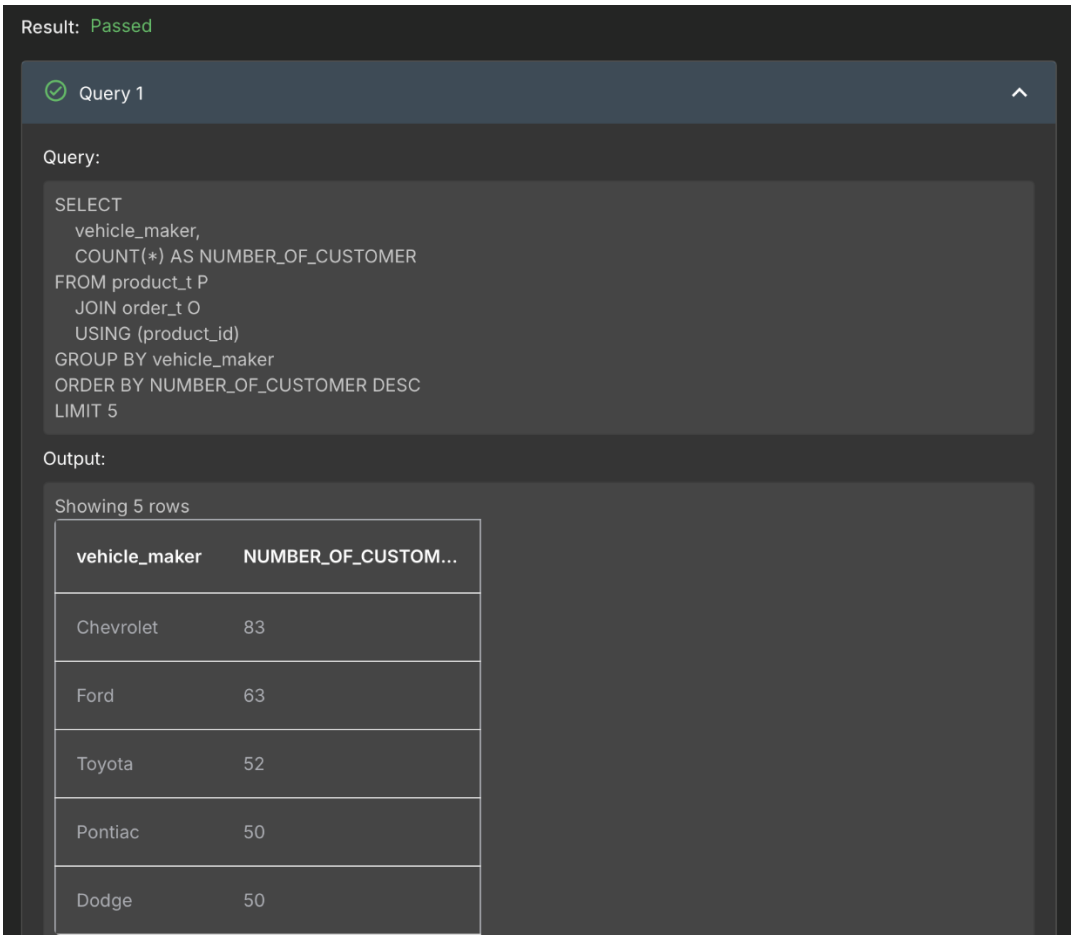
- Most Number of customers are residing in Texas (97) and California (97) followed by Florida, New York

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

Solution Query:

```
SELECT
    vehicle_maker,
    COUNT(*) AS NUMBER_OF_CUSTOMER
FROM product_t P
    JOIN order_t O
    USING (product_id)
GROUP BY vehicle_maker
ORDER BY NUMBER_OF_CUSTOMER DESC
LIMIT 5;
```

Output:



Observations and Insights:

- Chevrolet(83) is the most preferred vehicle By Customers followed by Ford(63), Toyota(52), Pontiac(50), Dodge(50).

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

Solution Query:

```
SELECT * FROM (
    SELECT
        state,
        vehicle_maker,
        COUNT(customer_id) AS TOTAL_CUSTOMER,
        RANK() OVER(PARTITION BY state ORDER BY COUNT(customer_id) DESC) AS
RNK
    FROM product_t P
        JOIN order_t O
        USING (product_id)
        JOIN customer_t C
        USING (customer_id)
    GROUP BY state, vehicle_maker
) AS PREF_VEHICLE
WHERE RNK = 1
ORDER BY TOTAL_CUSTOMER DESC;
```

Output:

Query 1

Query:

```
SELECT * FROM (
   SELECT
      state,
      vehicle_maker,
      COUNT(customer_id) AS TOTAL_CUSTOMER,
      RANK() OVER(PARTITION BY state ORDER BY COUNT(customer_id) DESC) AS RNK
   FROM product_t P
      JOIN order_t O
      USING (product_id)
      JOIN customer_t C
      USING (customer_id)
   GROUP BY state, vehicle_maker
) AS PREF_VEHICLE
WHERE RNK = 1
ORDER BY TOTAL_CUSTOMER DESC
```

Output:

Showing first 10 rows out of 143 rows

state	vehicle_maker	TOTAL_CUSTOMER	RNK
Texas	Chevrolet	9	1
Florida	Toyota	7	1
California	Nissan	6	1
California	Ford	6	1

Observations and Insights:

- Most Preferred vehicle In Texas is Chevrolet with highest customers count being 9.
- In Florida Toyota is most preferred vehicle with 7 customers.

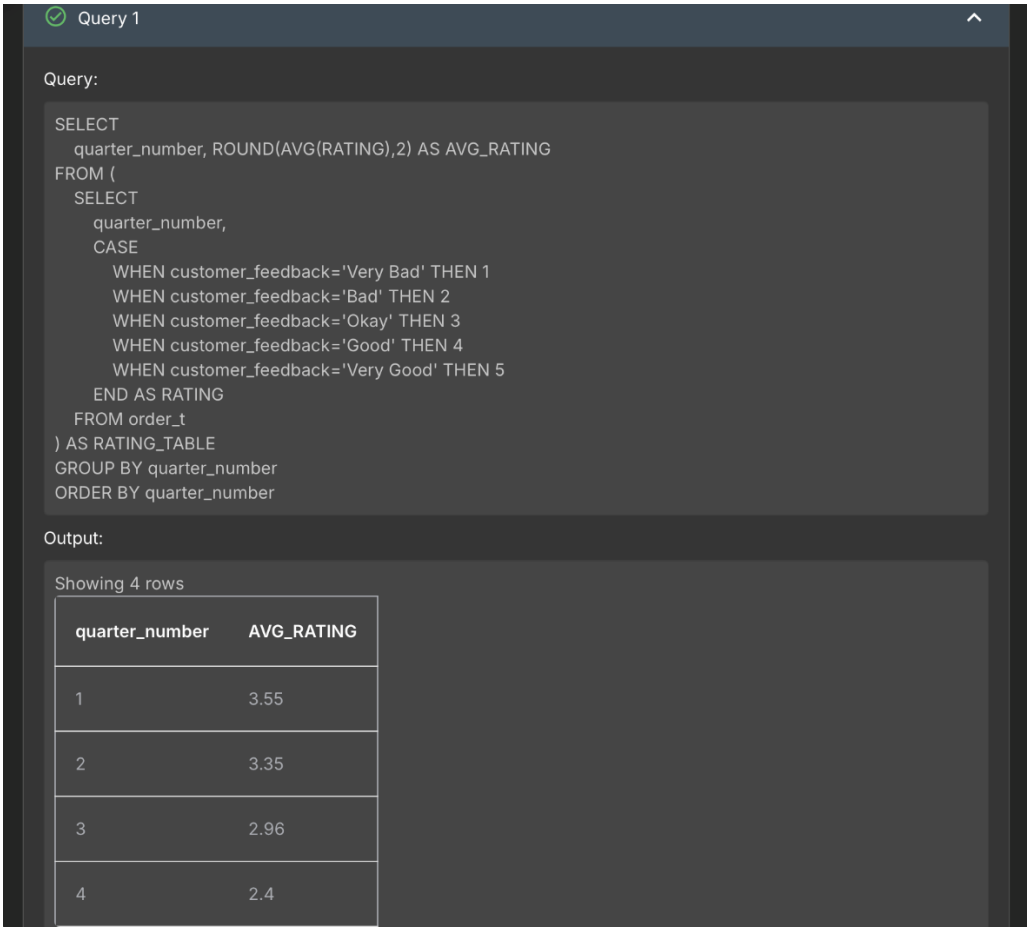
Question 4: Find the overall average rating given by the customers. What is the average rating in each quarter? Consider the following mapping for ratings: “Very Bad”: 1, “Bad”: 2, “Okay”: 3, “Good”: 4, “Very Good”: 5

Solution Query:

```
SELECT
    quarter_number, ROUND(AVG(RATING),2) AS AVG_RATING
FROM (
    SELECT
        quarter_number,
        CASE
            WHEN customer_feedback='Very Bad' THEN 1
            WHEN customer_feedback='Bad' THEN 2
            WHEN customer_feedback='Okay' THEN 3
            WHEN customer_feedback='Good' THEN 4
            WHEN customer_feedback='Very Good' THEN 5
```

```
        END AS RATING
    FROM order_t
) AS RATING_TABLE
GROUP BY quarter_number
ORDER BY quarter_number;
```

Output:



Observations and Insights:

- It is evident form the data that customer rating is decreasing with each quarter.
- While Quarter 1 is having Okay / Good Avg Customer rating 3.55, Quarter 4 is having Bad Avg Customer rating 2.4.

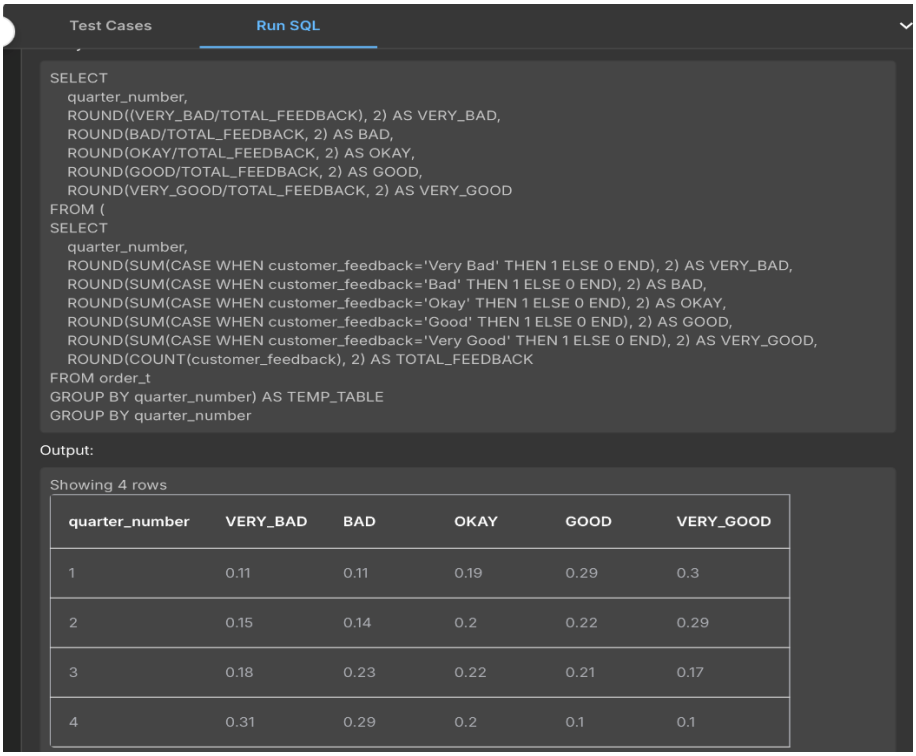
Question 5: Find the percentage distribution of feedback from the customers. Are customers getting more dissatisfied over time?

Solution Query:

```
SELECT
    quarter_number,
    ROUND((VERY_BAD/TOTAL_FEEDBACK), 2) AS VERY_BAD,
    ROUND(BAD/TOTAL_FEEDBACK, 2) AS BAD,
    ROUND(OKAY/TOTAL_FEEDBACK, 2) AS OKAY,
```

```
    ROUND(GOOD/TOTAL_FEEDBACK, 2) AS GOOD,
    ROUND(VERY_GOOD/TOTAL_FEEDBACK, 2) AS VERY_GOOD
FROM (
SELECT
    quarter_number,
    ROUND(SUM(CASE WHEN customer_feedback='Very Bad' THEN 1 ELSE 0 END), 2)
AS VERY_BAD,
    ROUND(SUM(CASE WHEN customer_feedback='Bad' THEN 1 ELSE 0 END), 2) AS
BAD,
    ROUND(SUM(CASE WHEN customer_feedback='Okay' THEN 1 ELSE 0 END), 2) AS
OKAY,
    ROUND(SUM(CASE WHEN customer_feedback='Good' THEN 1 ELSE 0 END), 2) AS
GOOD,
    ROUND(SUM(CASE WHEN customer_feedback='Very Good' THEN 1 ELSE 0 END), 2)
AS VERY_GOOD,
    ROUND(COUNT(customer_feedback), 2) AS TOTAL_FEEDBACK
FROM order_t
GROUP BY quarter_number) AS TEMP_TABLE
GROUP BY quarter_number;
```

Output:



Observations and Insights:

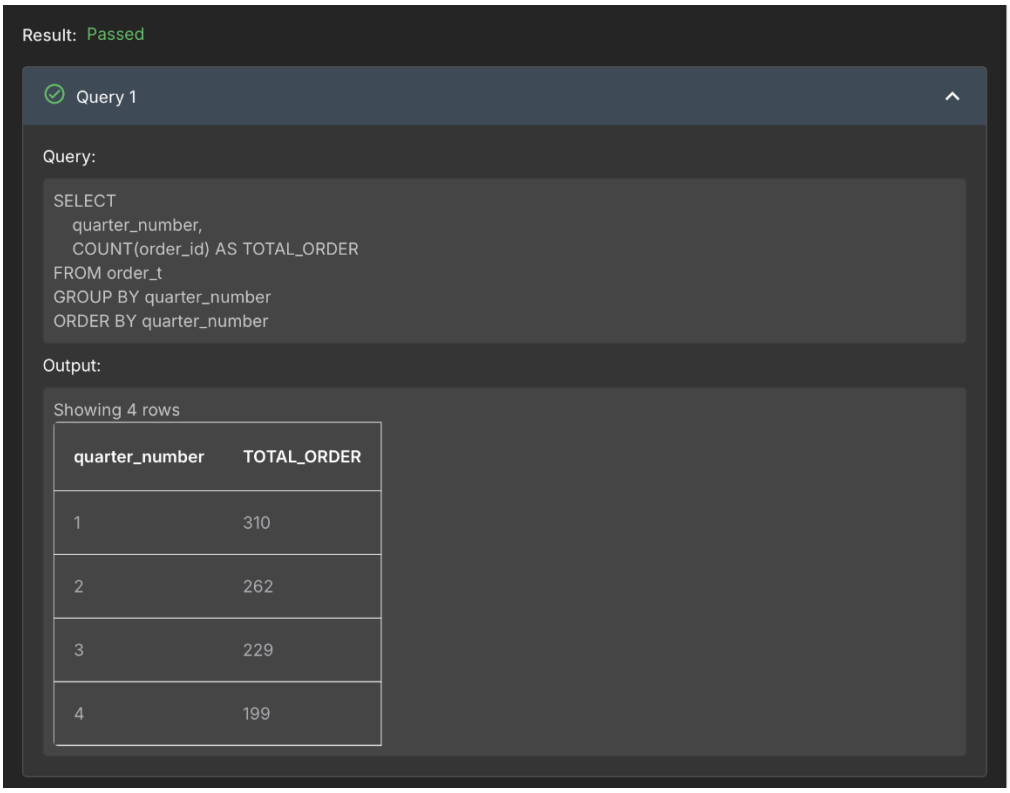
- The rating of Very Bad and Bad has increased from 0.11 to 0.31, 0.29 respectively, whereas Good and Verify Good rating decreased from 0.29, 0.3 to 0.1.
- From the output numbers, it shows that customers are getting more dissatisfied over time.

Question 6: What is the trend of the number of orders by quarter?

Solution Query:

```
SELECT
    quarter_number,
    COUNT(order_id) AS TOTAL_ORDER
FROM order_t
GROUP BY quarter_number
ORDER BY quarter_number;
```

Output:



Observations and Insights:

- Number of orders decreased over time from Q1 to Q4.
- In Q1, the number of orders was 310 where as in Q4 it reduced to 199, resulting in 35.81% decline.
- There is a negative trend in the number of orders per quarter.

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

Solution Query:

```
SELECT
    quarter_number,
    REVENUE,
    ROUND(LAG(REVENUE) OVER(ORDER BY quarter_number), 2) AS
PREVIOUS_REVENUE,
```

```
    ROUND((REVENUE - LAG(REVENUE) OVER(ORDER BY quarter_number))/
LAG(REVENUE) OVER(ORDER BY quarter_number), 2) AS QoQ
FROM (
    SELECT
        quarter_number,
        ROUND(SUM(quantity * (vehicle_price-((discount /
100)*vehicle_price))), 2) AS REVENUE
    FROM order_t
    GROUP BY quarter_number
) AS QUATERLY_REVENUE
ORDER BY quarter_number;
```

Output:

Result: Passed

Query 1

Query:

```
SELECT
   quarter_number,
   REVENUE,
   ROUND(LAG(REVENUE) OVER(ORDER BY quarter_number), 2) AS PREVIOUS_REVENUE,
   ROUND((REVENUE - LAG(REVENUE) OVER(ORDER BY quarter_number))/LAG(REVENUE) OVER(ORDER BY
quarter_number), 2) AS QoQ
FROM (
   SELECT
      quarter_number,
      ROUND(SUM(quantity * (vehicle_price-((discount / 100)*vehicle_price))), 2) AS REVENUE
   FROM order_t
   GROUP BY quarter_number
) AS QUATERLY_REVENUE
ORDER BY quarter_number
```

Output:

Showing 4 rows

quarter_number	REVENUE	PREVIOUS_REVENUE	QoQ
1	39421580.16		
2	32715830.34	39421580.16	-0.17
3	29229896.19	32715830.34	-0.11
4	23346779.63	29229896.19	-0.2

Observations and Insights:

- It shows that there is decrease in revenue when compared with previous quarter revenue.
- When we compare per quarter revenue decline, it has decreased -17%, -11%, -20% per quarter

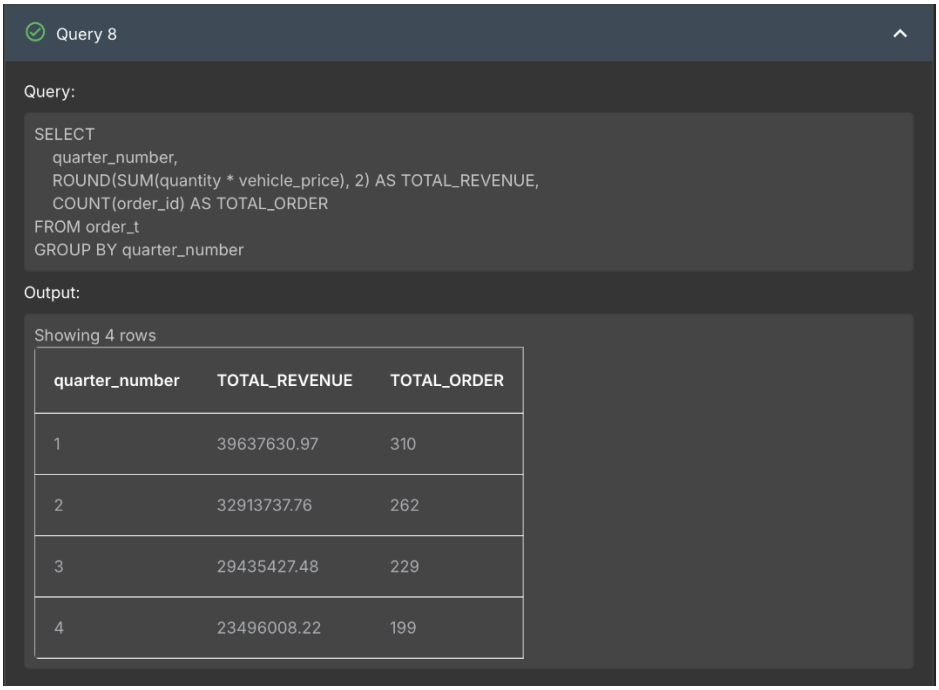
Question 8: What is the trend of net revenue and orders by quarters?

Solution Query:

```
SELECT
```

```
    quarter_number,
    ROUND(SUM(quantity * vehicle_price), 2) AS TOTAL_REVENUE,
    COUNT(order_id) AS TOTAL_ORDER
FROM order_t
GROUP BY quarter_number;
```

Output:



Observations and Insights:

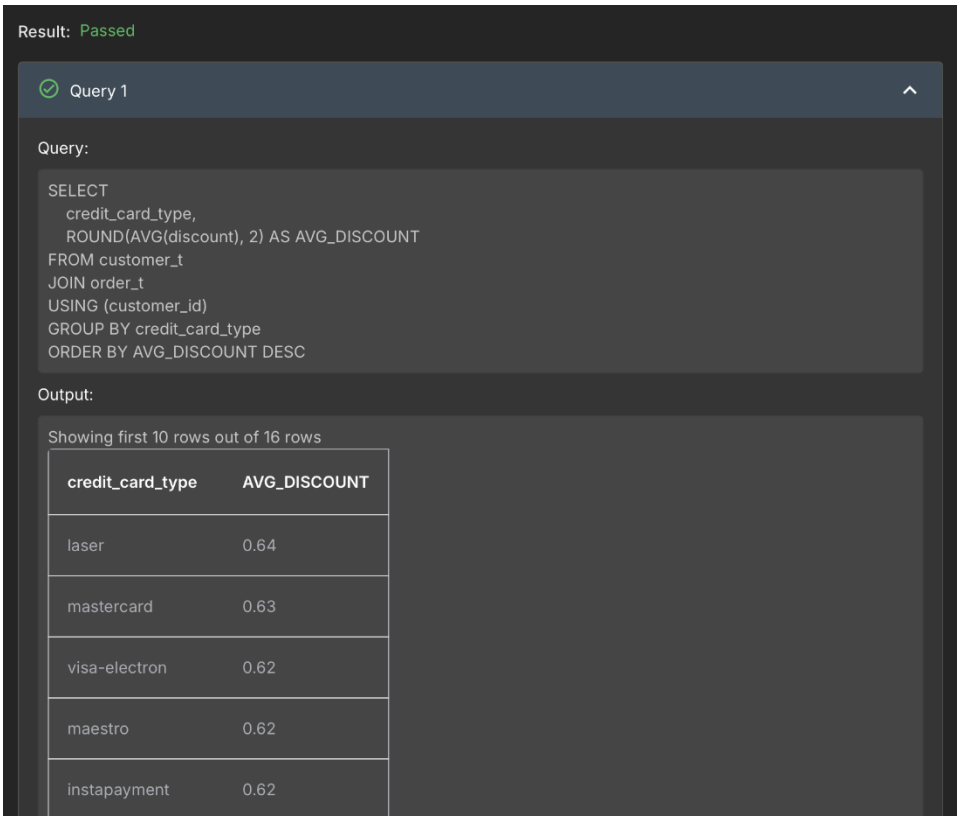
- Declining trend of revenue and total orders by quarter is observed from the output.
- Revenue decreased from 39.64M to 23.49M, showing a 40% reduction.
- It also shows that total order has decreased from 310 to 199.

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

Solution Query:

```
SELECT
    credit_card_type,
    ROUND(AVG(discount), 2) AS AVG_DISCOUNT
FROM customer_t
JOIN order_t
USING (customer_id)
GROUP BY credit_card_type
ORDER BY AVG_DISCOUNT DESC;
```

Output:



Observations and Insights:

- Highest discount is offered by laser 64% followed by MasterCard that is 63%.

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

Solution Query:

```
SELECT
    quarter_number,
    ROUND(AVG(JULIANDAY(DATE(ship_date)) - JULIANDAY(DATE(order_date))), 0)
AS AVG_TIME
FROM order_t
GROUP BY quarter_number
ORDER BY quarter_number;
```

Output:

Query 10

Query:

```
SELECT
   quarter_number,
   ROUND(AVG(JULIANDAY(DATE(ship_date)) - JULIANDAY(DATE(order_date))), 0) AS AVG_TIME
FROM order_t
GROUP BY quarter_number
ORDER BY quarter_number
```

Output:

Showing 4 rows

quarter_number	AVG_TIME
1	57
2	71
3	118
4	174

Observations and Insights:

- There is a delay in Average shipping time from Q1 to Q4.
- In Quater 1, Average shipping time was 57 days where in Q4 its 174 days, which is 117 Days more.

Business Metrics Overview

Total Revenue	Total Orders	Total Customers	Average Rating
125.4M	1000	994	3.14
Last Quarter Revenue	**Last quarter Orders**	**Average Days to Ship**	**% Good Feedback**
23.4M	199	98	22%

Business Recommendations

- There should be more focus on how to improve customer experience. A dedicated team should be selected who can conduct surveys to gather customer pain points and identify root cause.
- To retain customers and create loyalty base, personalized promotions should be sent out to targeted customers.
- On low performing areas of sale, a marketing campaign can be organized in social media to target customers and make them aware of our new promotions.
- To improve shipping time for orders, there should be focus on better logistics. This can be done by partnering with reliable shipping companies who can deliver order by the given date.