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:

```
select
      ct.state,
      count(distinct ot.order_id) as customer_count
from customer_t ct
      join order_t ot on ct.customer_id = ot.customer_id
group by ct.state
order by customer_count desc;
```

Output:

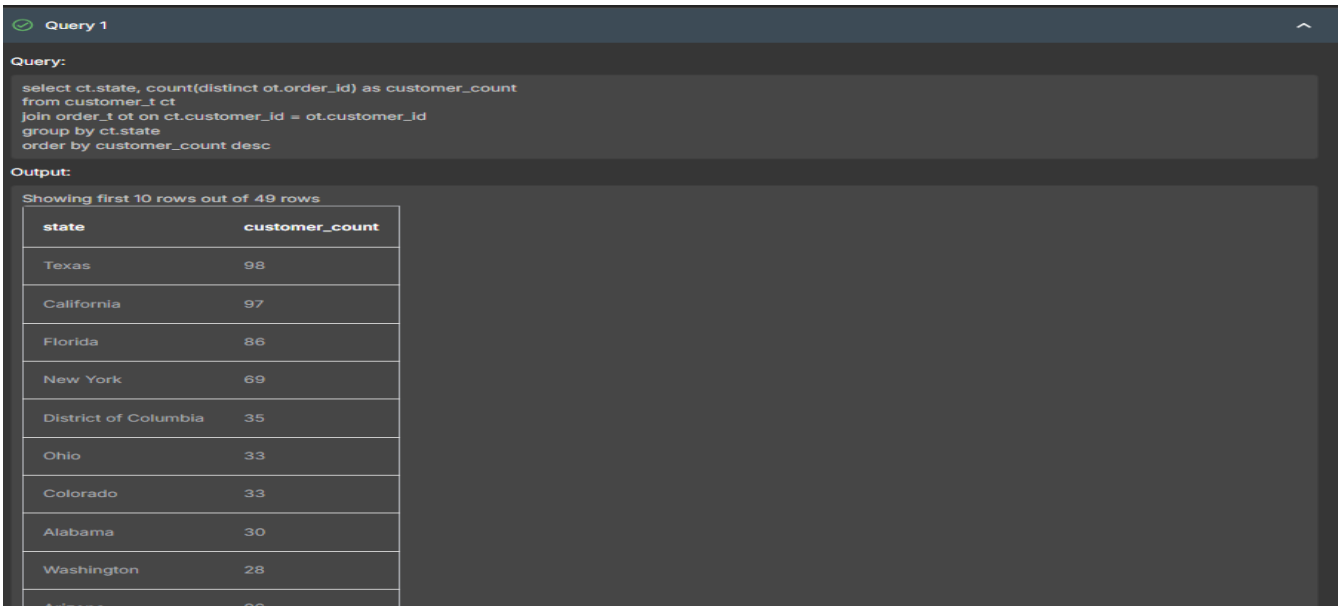

Figure 1

Observations and Insights:

- The output contains a total of 49 rows, representing the number of customers who placed orders, distributed across different states.

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

Solution Query:

```
select
      pt.vehicle_maker,
      count(ot.order_id) as total_orders
from product_t pt
      join order_t ot on pt.product_id = ot.product_id
group by pt.vehicle_maker
order by total_orders desc
      limit 5;
```

Output:

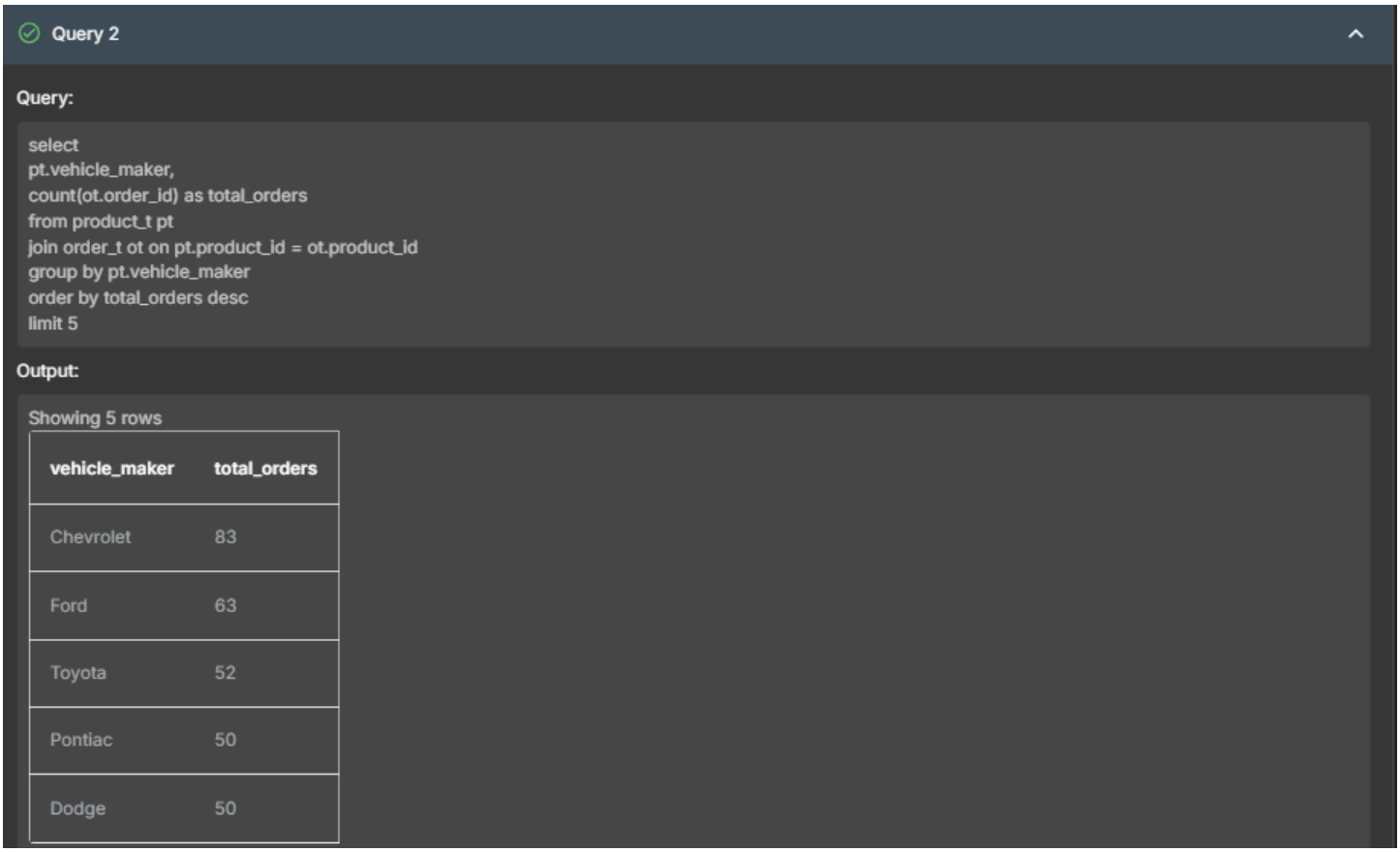

Figure 2

Observations and Insights:

- The top 5 vehicle makers were identified based on the number of orders placed by customers.
- As shown in Figure 2, the top 5 vehicle makers are:
    1. Chevrolet with 83 orders.
    2. Ford with 63 orders.

3. Toyota with 52 orders.
4. Pontiac with 50 orders.
5. Dodge with 50 orders.

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

Solution Query:

```
select
      distinct ct.state,
      pt.vehicle_maker,
      count(ot.order_id) as total_orders,
      rank() over (partition by ct.state
                  order by count(ot.order_id)desc) as rnk
from order_t ot
    join product_t pt using (product_id)
    join customer_t ct using (customer_id)
group by ct.state, pt.vehicle_maker
order by total_orders desc
      limit 10;
```

Output:

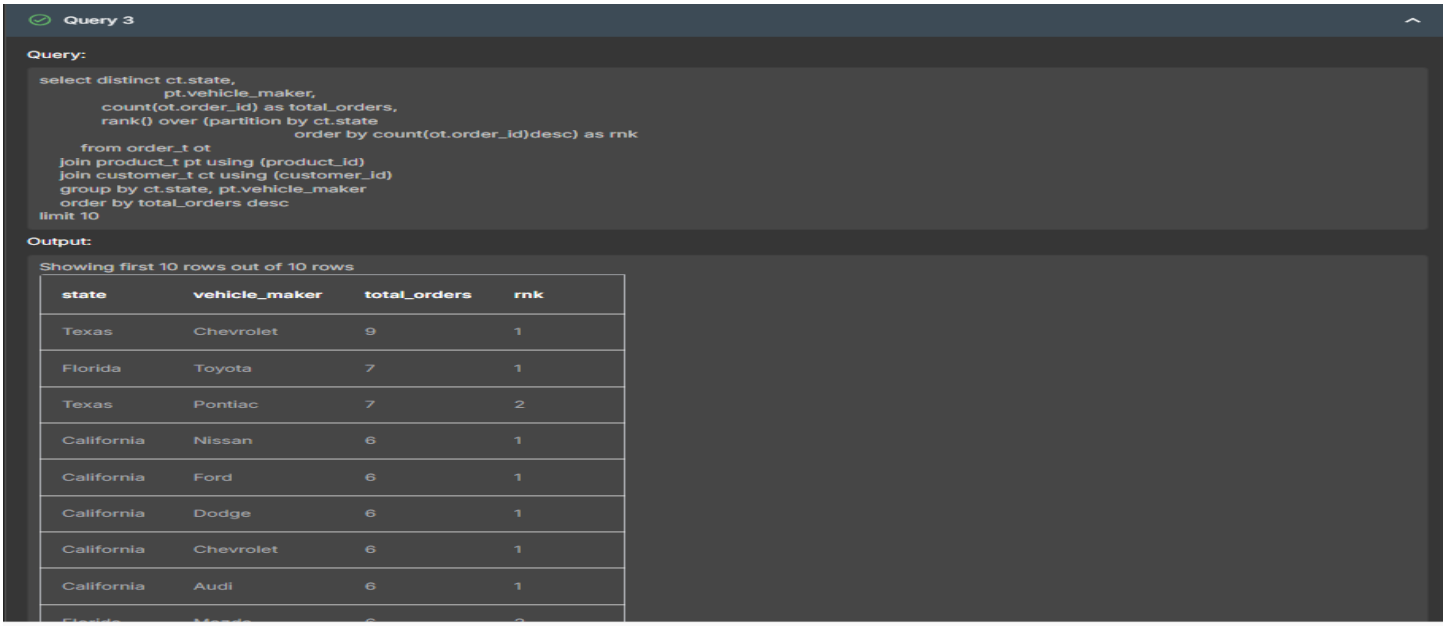

Figure 3

Observations and Insights:

- The output was presented with the respective rankings.
- As observed in Figure 3, Chevrolet is the most preferred vehicle maker, with a total of 9 orders from customers. Compared to other vehicle makers, Chevrolet was the top choice in the state of Texas.

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
    'Overall' AS quarter_number,
    AVG(
        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 average_rating
FROM order_t

UNION ALL

SELECT
    quarter_number,
    AVG(
        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 average_rating
FROM order_t
GROUP BY quarter_number;
```

Output:

Query 4

Query:

```
SELECT
    'Overall' AS quarter_number,
    AVG(
        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 average_rating
FROM order_t

UNION ALL

SELECT
    quarter_number,
    AVG(
        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 average_rating
FROM order_t
GROUP BY quarter_number
```

Output:

Showing 5 rows

quarter_number	average_rating
Overall	3.135
1	3.554838709677419
2	3.354961832061069
3	2.9563318777292578
4	2.3969849246231156

Figure 4

Observations and Insights:

- The overall average rating is 3.135.
- The average ratings for each quarter are as follows:
    1. Quarter 1: 3.55
    2. Quarter 2: 3.35
    3. Quarter 3: 2.95
    4. Quarter 4: 2.39

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

Solution Query:

```
select
    quarter_number,
    count(case when customer_feedback = 'very bad' then 1 end) * 100.0 / count(*)
as very_bad_percentage,
    count(case when customer_feedback = 'bad' then 1 end) * 100.0 / count(*) as
bad_percentage,
    count(case when customer_feedback = 'okay' then 1 end) * 100.0 / count(*) as
okay_percentage,
    count(case when customer_feedback = 'good' then 1 end) * 100.0 / count(*) as
good_percentage,
    count(case when customer_feedback = 'very good' then 1 end) * 100.0 / count(*)
as very_good_percentage,
    count(case when customer_feedback in ('bad', 'very bad') then 1 end) * 100.0 /
count(*) as dissatisfaction_percentage
from order_t
group by quarter_number
order by quarter_number;
```

Output:

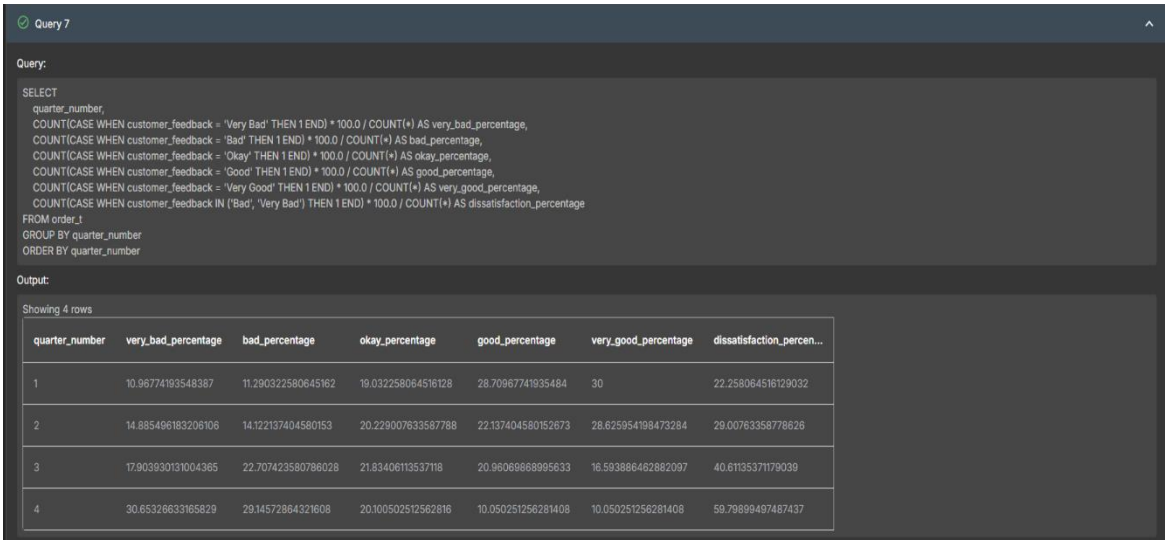

Figure 5

Observations and Insights:

- The percentage distribution of customer feedback is illustrated in Figure 5.
- A significant rise in dissatisfaction is observed in the last quarter, reaching approximately 60%, indicating an increase in dissatisfaction towards the end of the year.In contrast, the first quarter recorded the lowest dissatisfaction rate at 22.25%.

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

Solution Query:

```
select
      quarter_number,
      count(*) as total_orders,
      count(*) * 100.0 / sum(count(*)) over()) as trend_perc
from order_t
      group by quarter_number
      order by quarter_number asc;
```

Output:

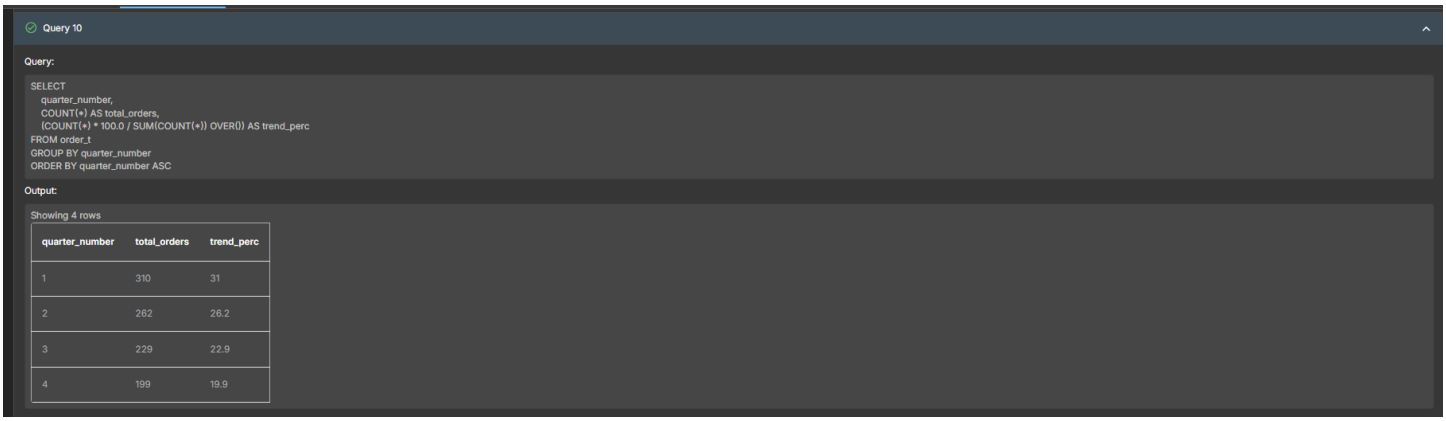

Figure 6

Observations and Insights:

- As per the figure **Error! Reference source not found.**, the trend percentage of orders took place in first 3 months are ~ 31% which is higher
- Therefore, the trend is observed from higher percentage to low, i.e., order decreased towards the end 31% to 20%
- ~10% decrease orders can be observed.

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

Solution Query:

```
with quarterlyrevenue as (
    select
        quarter_number,
        sum(vehicle_price - discount) as net_revenue
    from order_t
    group by quarter_number),

revenuechange as (
    select
        quarter_number,
        net_revenue,
        lag(net_revenue) over (order by quarter_number) as prev_revenue,
        round(
            ((net_revenue - lag(net_revenue) over (order by quarter_number))
            / lag(net_revenue) over (order by quarter_number)) * 100, 2
        ) as qoq_change
    from quarterlyrevenue
)
select * from revenuechange;
```

Output:

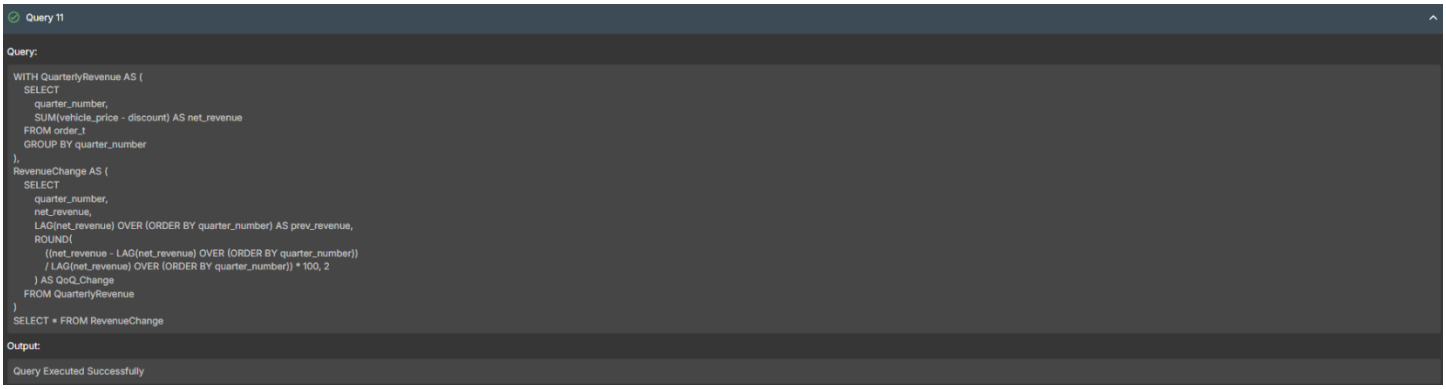

Figure 7

Result Grid | Filter Rows: | Export: | Wrap C

quarter_number	net_revenue	prev_revenue	QoQ_Change
1	26519030.71	NULL	NULL
2	21595716.63	26519030.71	-18.57
3	19719757.68	21595716.63	-8.69
4	15279883.03	19719757.68	-22.51

Figure 8

Observations and Insights:

- Figure 8 shows the net revenue generated quarter -over -quarter % by the company.
- Positive QoQ change shows that the revenue increased and negative QoQ change shows that the Revenue declined. Here, the observation shows decline in the QoQ change as the values are negative.

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

Solution Query:

```
select
    quarter_number,
    count(*) as total_orders,
    sum(vehicle_price - discount) as net_revenue
from order_t
group by  quarter_number
order by  quarter_number asc;
```

Output:

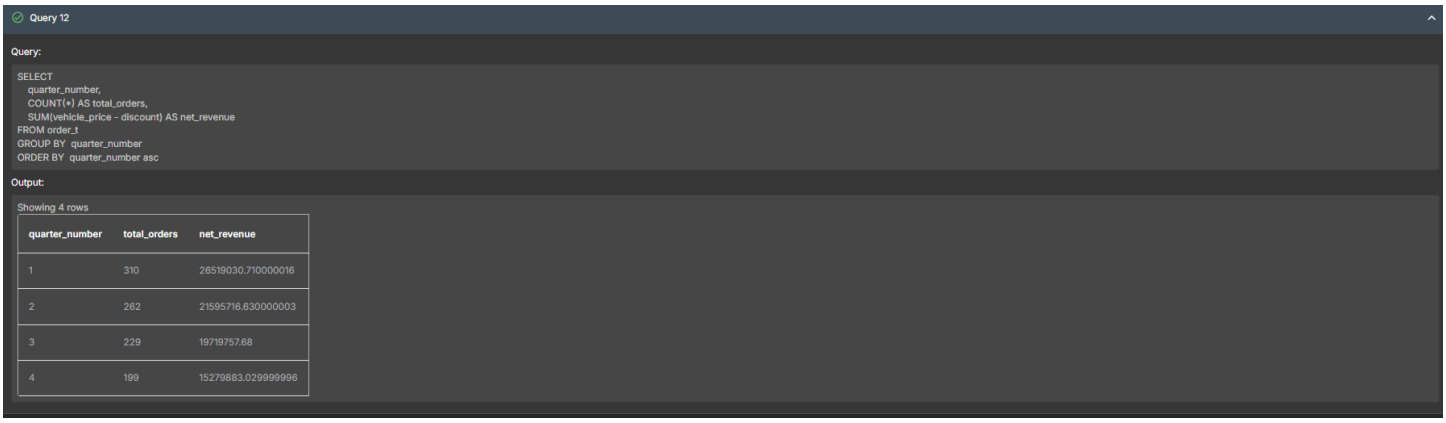

Figure 9

Observations and Insights:

- The net revenue in the first quarter was 26,519,030.71, which declined in the second quarter to 21,595,716.63, followed by 19,719,757.68 in the third quarter and 15,279,883.03 in the fourth quarter.

- This indicates that the first quarter had the highest net revenue, with a gradual decline towards the end of the year.

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

Solution Query:

```
select
      ct.credit_card_type as credit_cards,
    avg(ot.discount) as average_discount_offered
from customer_t ct
      join order_t ot using(customer_id)
group by ct.credit_card_type;
```

Output:

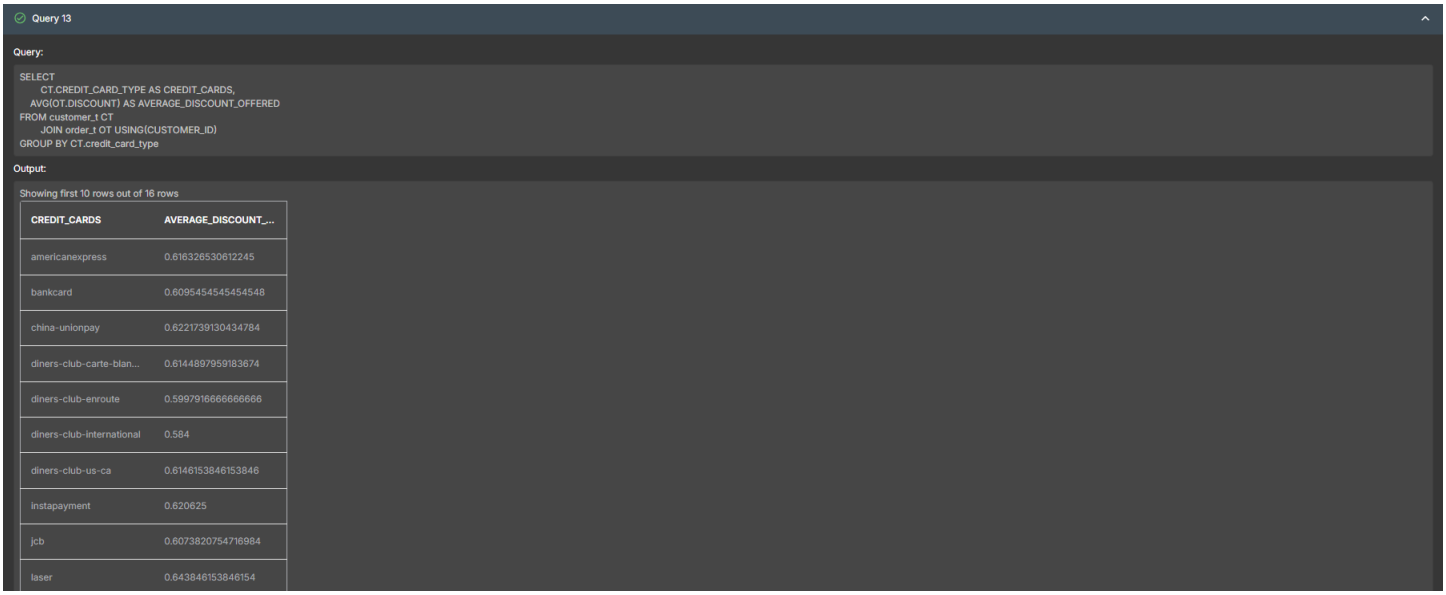

Figure 10

Observations and Insights:

- 16 rows returned with the average discount offered to respective credit cards in figure 10.
- Credit cards like '*solo*' and '*diners-club-international*' offers low discounts comparing to other credit cards.
- Most of the credit cards offers discount above 60.0%

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

Solution Query:

```
select
    quarter_number,
    avg(julianday(ship_date) - julianday(order_date)) as average_time
from order_t
    group by quarter_number
    order by average_time asc;
```

Output:

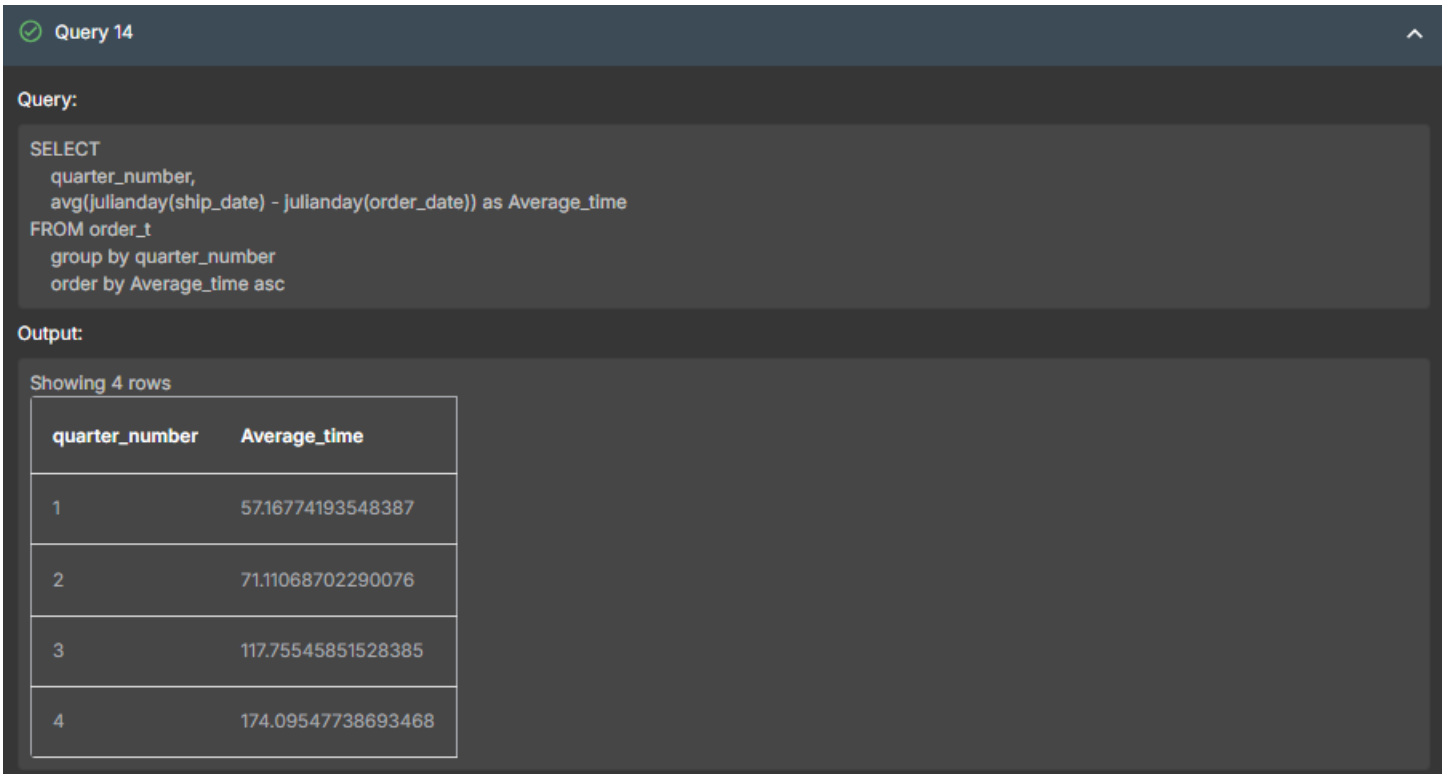

Figure 11

Observations and Insights:

- The average time taken for the shipping of order once the order placed for each quarter is given in figure11.
- Maximum average time was taken for the quarter 4 i.e., 175 days
- Minimum average time was taken for the quarter 1 i.e., 58 days
- The average time taken to ship the order is increased from 1st quarter to 4th quarter.

Business Metrics Overview

Total Revenue	Total Orders	Total Customers	Average Rating
125482191.37	1000.0	994.0	3.1350
Last Quarter Revenue	**Last quarter Orders**	**Average Days to Ship**	**% Good Feedback**
23495881.27	199	~98 days	44.1%

Business Recommendations

- Total revenue of $125.48M indicates overall sales performance is strong. Whereas last quarter revenue is $23.49 M which might shows the seasonal trends or declining sales towards the end. The solution for this is by boosting seasonal marketing campaigns or offering subscription-based models for recurring purchases.

- There are 1000.0 orders which indicates healthy demand in the sales, but based on the trends, sales can improve much more. Last quarter orders are 199 suggests the drop of order volume compare to overall numbers.

- Total customers are 994.0 which is near to total orders of 1000. The values are close which means almost all customers did the purchase implies that there is a potential to improve the purchase and repurchase chances. Utilize the customers by incorporating the loyalty programs and special discounts, personalized email campaigns for the repurchases

- Average customer experience is 3.135 which indicates improvement is required. Therefore, improve the customer satisfaction and Ratings by collecting detailed feedback, give discounts/vouchers for the dissatisfied customers, etc.

- One of the major reasons for the customer dissatisfaction might be the average days to ship the orders once placed. 98 days are extremely higher number of days for the customers to receive their orders. Reduce the shipping day duration by partnering with fastest delivery services/ express delivery system.

- Less than half of customers rated their experience as "Good" or "Very Good". It can be overcome by offering post-purchase support & engagement to the customers or train customer service teams to handle negative feedback proactively.