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
    'Total' AS state,
    COUNT(DISTINCT o.customer_id) AS customer_count
FROM
    customer_t c
JOIN
    order_t o ON c.customer_id = o.customer_id

UNION ALL

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

Output:

Result: Passed

Query 1

Query:

```
SELECT
   'Total' AS state,
   COUNT(DISTINCT o.customer_id) AS customer_count
FROM
   customer_t c
JOIN
   order_t o ON c.customer_id = o.customer_id

UNION ALL

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

Output:

Showing first 10 rows out of 50 rows

state	customer_count
Total	994
Alabama	29
Alaska	10
Arizona	26
Arkansas	6
California	97
Colorado	33
Connecticut	22
Delaware	6
District of Columbia	35

Observations and Insights:

- California has a high number of customers, indicating it could be a profitable market for the company.
- States with low customer concentration such as Delaware or Arkansas could be targeted through marketing campaigns to increase sales in those areas.

- With 994 total unique customers, we can analyze this dataset to provide insights into customer preferences and behavior.

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

Solution Query:

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

Output:

Result: Passed

Query 1

Query:

```
SELECT
   p.vehicle_maker,
   COUNT(o.product_id) AS total_orders
FROM
   order_t o
JOIN
   product_t p ON o.product_id = p.product_id
GROUP BY
   p.vehicle_maker
ORDER BY
   total_orders DESC
LIMIT 5
```

Output:

Showing 5 rows

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

Observations and Insights:

- Chevrolet has the highest number of total orders indicating strong customer preference.
- Toyota, a foreign brand, ranks higher in sales than domestic makers Pontiac and Dodge.
- The closeness in Toyota, Pontiac, and Dodge shows that customers may be open to exploring other brands besides the dominant makers (Ford and Chevrolet).

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

Solution Query:

```
SELECT
    state,
    vehicle_maker,
    total_orders
FROM (
    SELECT
        c.state,
        p.vehicle_maker,
        COUNT(o.product_id) AS total_orders,
        ROW_NUMBER() OVER (PARTITION BY c.state ORDER BY COUNT(o.product_id) DESC) AS rank
    FROM
        customer_t c
    JOIN
        order_t o ON c.customer_id = o.customer_id
    JOIN
        product_t p ON o.product_id = p.product_id
    GROUP BY
        c.state, p.vehicle_maker
) AS ranked_results
WHERE rank = 1;
```

Output:

Result: Passed

Query 1

Query:

```
SELECT
    state,
    vehicle_maker,
    total_orders
FROM (
    SELECT
        c.state,
        p.vehicle_maker,
        COUNT(o.product_id) AS total_orders,
        ROW_NUMBER() OVER (PARTITION BY c.state ORDER BY COUNT(o.product_id) DESC) AS rank
    FROM
        customer_t c
    JOIN
        order_t o ON c.customer_id = o.customer_id
    JOIN
        product_t p ON o.product_id = p.product_id
    GROUP BY
        c.state, p.vehicle_maker
) AS ranked_results
WHERE rank = 1
```

Output:

Showing first 10 rows out of 49 rows

state	vehicle_maker	total_orders
Alabama	Dodge	5
Alaska	Chevrolet	2
Arizona	Pontiac	3
Arkansas	Volkswagen	1
California	Nissan	6
Colorado	Chevrolet	5
Connecticut	Volvo	2
Delaware	Mitsubishi	2
District of Columbia	Chevrolet	4
Florida	Toyota	7

Observations and Insights:

- Chevrolet is the preferred vehicle maker in Alaska, Colorado, and District of Columbia.
- Florida accounts for approximately 13.46% of Toyotas total sales.
- Nissan is the most preferred brand in one of the most competitive markets (California).

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:

```
-- Overall average rating
SELECT
   'Overall' AS category,
   CASE
      WHEN avg_rating <= 1.5 THEN 'Very Bad'
      WHEN avg_rating <= 2.5 THEN 'Bad'
      WHEN avg_rating <= 3.5 THEN 'Okay'
      WHEN avg_rating <= 4.5 THEN 'Good'
      ELSE 'Very Good'
   END AS average_rating
FROM (
   SELECT
      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 avg_rating
   FROM order_t
)

UNION ALL

-- Average rating per quarter
SELECT
   'Quarter ' || quarter_number AS category,
```

```
    CASE
        WHEN avg_rating <= 1.5 THEN 'Very Bad'
        WHEN avg_rating <= 2.5 THEN 'Bad'
        WHEN avg_rating <= 3.5 THEN 'Okay'
        WHEN avg_rating <= 4.5 THEN 'Good'
        ELSE 'Very Good'
    END AS average_rating
FROM (
    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 avg_rating
    FROM order_t
    GROUP BY quarter_number
);
```

Output:

Result: Passed

Query 1

Query:

```
SELECT
   'Overall' AS category,
   CASE
      WHEN avg_rating <= 1.5 THEN 'Very Bad'
      WHEN avg_rating <= 2.5 THEN 'Bad'
      WHEN avg_rating <= 3.5 THEN 'Okay'
      WHEN avg_rating <= 4.5 THEN 'Good'
      ELSE 'Very Good'
   END AS average_rating
FROM (
   SELECT
      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 avg_rating
   FROM order_t
)

UNION ALL
```

Output:

Showing 5 rows

category	average_rating
Overall	Okay
Quarter 1	Good
Quarter 2	Okay
Quarter 3	Okay
Quarter 4	Bad

Observations and Insights:

- The year started strong in terms of customer ratings and trended downwards by the end of the year suggesting customer dissatisfaction/operational issues over time.
- Immediate attention to customer ratings, especially in Q4, should be analyzed by the company.
- Similar ratings in Q2 and Q3 suggest stagnation (neither improvement nor decline) in the customer experience.

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

Solution Query:

```
-- Percentage distribution of feedback per quarter
SELECT
    quarter_number,
    customer_feedback,
    ROUND((COUNT(*) * 100.0 / SUM(COUNT(*)) OVER (PARTITION BY quarter_number)), 2) AS
percentage_distribution
FROM
    order_t
GROUP BY
    quarter_number, customer_feedback
ORDER BY
    quarter_number, percentage_distribution DESC;

-- Dissatisfaction trend over time
SELECT
    quarter_number,
    ROUND(
        (SUM(
            CASE
                WHEN customer_feedback = 'Very Bad' OR customer_feedback = 'Bad' THEN 1
                ELSE 0
            END) * 100.0 / COUNT(*)),
        2
    ) AS dissatisfaction_percentage
FROM
    order_t
GROUP BY
    quarter_number
ORDER BY
    quarter_number;
```

Output (1):

Result: Passed

Query 1

Query:

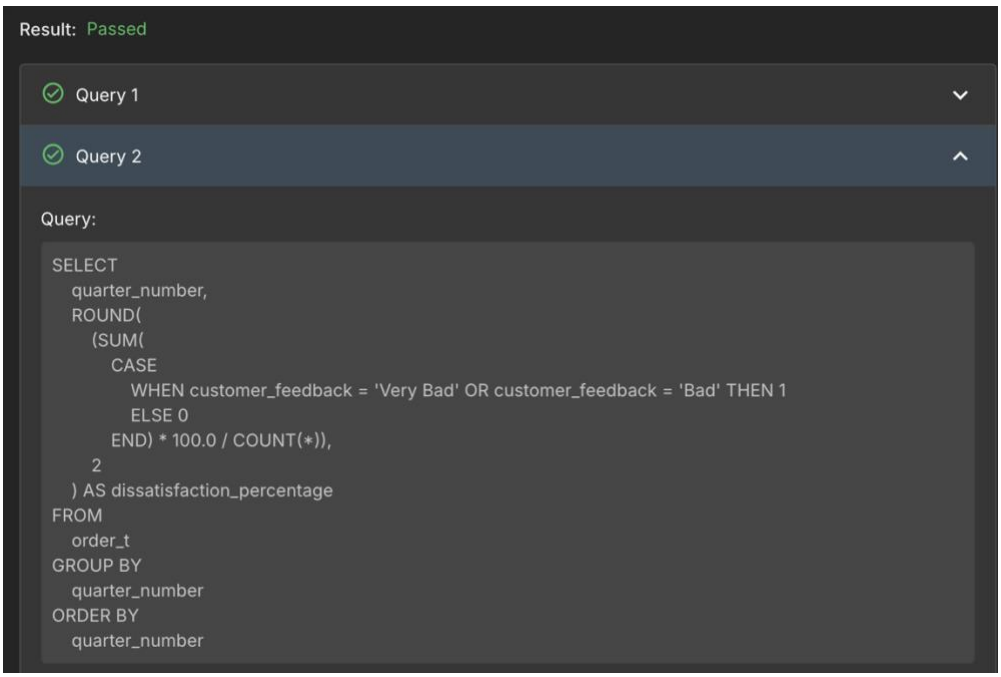
```
SELECT
   quarter_number,
   customer_feedback,
   ROUND((COUNT(*) * 100.0 / SUM(COUNT(*)) OVER (PARTITION BY quarter_number)), 2) AS
percentage_distribution
FROM
   order_t
GROUP BY
   quarter_number, customer_feedback
ORDER BY
   quarter_number, percentage_distribution DESC
```

Output:

Showing first 10 rows out of 20 rows

quarter_number	customer_feedback	percentage_distribution
1	Very Good	30
1	Good	28.71
1	Okay	19.03
1	Bad	11.29
1	Very Bad	10.97
2	Very Good	28.63
2	Good	22.14
2	Okay	20.23
2	Very Bad	14.89
2	Bad	14.12

Output (2):



Observations and Insights:

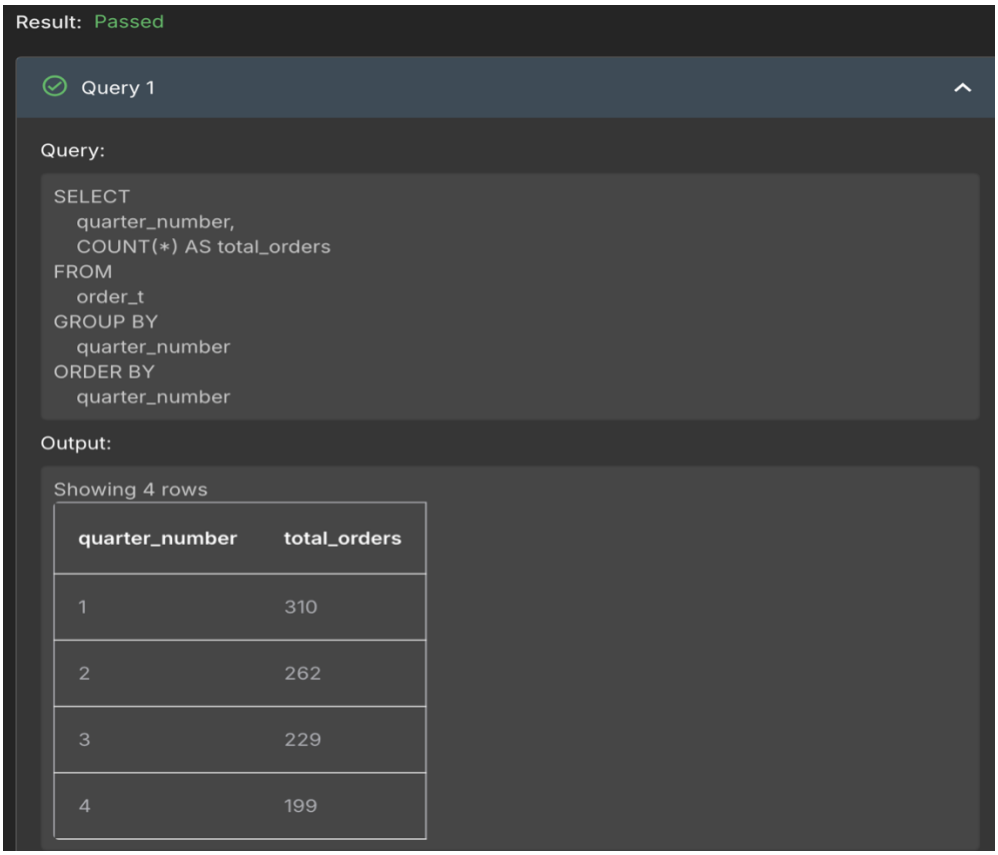
- The increasing trend in customer dissatisfaction suggests operational issues that are compounding over time. Investigating Q3 and Q4 could help identify the cause.
- Over half of the customers in Q4 were dissatisfied with the company.
- Customer dissatisfaction drastically increased by ~47% between Q3 and Q4.

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

Solution Query:

```
SELECT
    quarter_number,
    COUNT(*) AS total_orders
FROM
    order_t
GROUP BY
    quarter_number
ORDER BY
    quarter_number;
```

Output:



Observations and Insights:

- Total orders have decreased quarter over quarter.
- The largest decline in orders happened between Q1 and Q2, with a 15.5% decrease in sales.
- The consistent decline over quarters raises questions about customer retention strategies and is congruent with customer dissatisfaction data.

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

Solution Query:

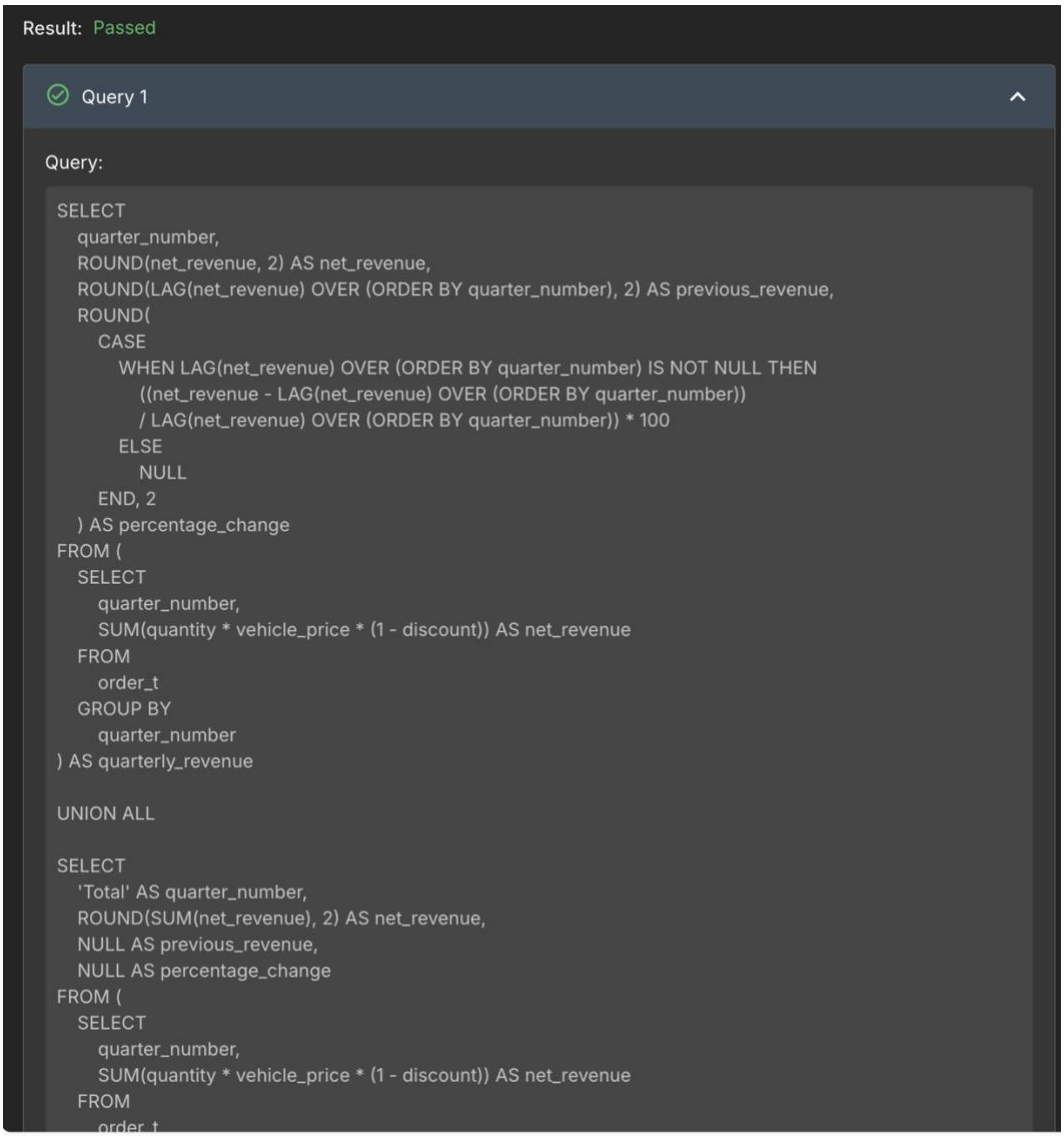
```
SELECT
    quarter_number,
    ROUND(net_revenue, 2) AS net_revenue,
    ROUND(LAG(net_revenue) OVER (ORDER BY quarter_number), 2) AS previous_revenue,
    ROUND(
        CASE
            WHEN LAG(net_revenue) OVER (ORDER BY quarter_number) IS NOT NULL THEN
                ((net_revenue - LAG(net_revenue) OVER (ORDER BY quarter_number))
                / LAG(net_revenue) OVER (ORDER BY quarter_number)) * 100
            ELSE
                NULL
        END, 2
    ) AS percentage_change
FROM (
    SELECT
        quarter_number,
        SUM(quantity * vehicle_price * (1 - discount)) AS net_revenue
    FROM
        order_t
    GROUP BY
        quarter_number
) AS quarterly_revenue

UNION ALL

SELECT
    'Total' AS quarter_number,
    ROUND(SUM(net_revenue), 2) AS net_revenue,
    NULL AS previous_revenue,
    NULL AS percentage_change
FROM (
    SELECT
```

```
    quarter_number,
    SUM(quantity * vehicle_price * (1 - discount)) AS net_revenue
  FROM
    order_t
  GROUP BY
    quarter_number
) AS quarterly_revenue;
```

Output:



Output:

Showing 5 rows

quarter_number	net_revenue	previous_revenue	percentage_change
1	18032549.9		
2	13122995.76	18032549.9	-27.23
3	8882298.84	13122995.76	-32.32
4	8573149.28	8882298.84	-3.48
Total	48610993.78		

Observations and Insights:

- The net revenue from Q1 to Q2 drops 27.23%, suggesting market changes and/or reduced sales.
- Net revenue consistently declines quarter over quarter, but not as drastically between Q3 and Q4.
- The total net revenue for the year is $48,610,993.78, with Q1 contributing the largest share. Over 37% of the total revenue was generated in Q1 alone.

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

Solution Query:

```
SELECT
    quarter_number,
    ROUND(SUM(quantity * vehicle_price * (1 - discount)), 2) AS net_revenue,
    COUNT(*) AS total_orders
FROM
    order_t
GROUP BY
    quarter_number
ORDER BY
    quarter_number;
```

Output:

Query 1

Query:

```
SELECT
   quarter_number,
   ROUND(SUM(quantity * vehicle_price * (1 - discount)), 2) AS net_revenue,
   COUNT(*) AS total_orders
FROM
   order_t
GROUP BY
   quarter_number
ORDER BY
   quarter_number
```

Output:

Showing 4 rows

quarter_number	net_revenue	total_orders
1	18032549.9	310
2	13122995.76	262
3	8882298.84	229
4	8573149.28	199

Observations and Insights:

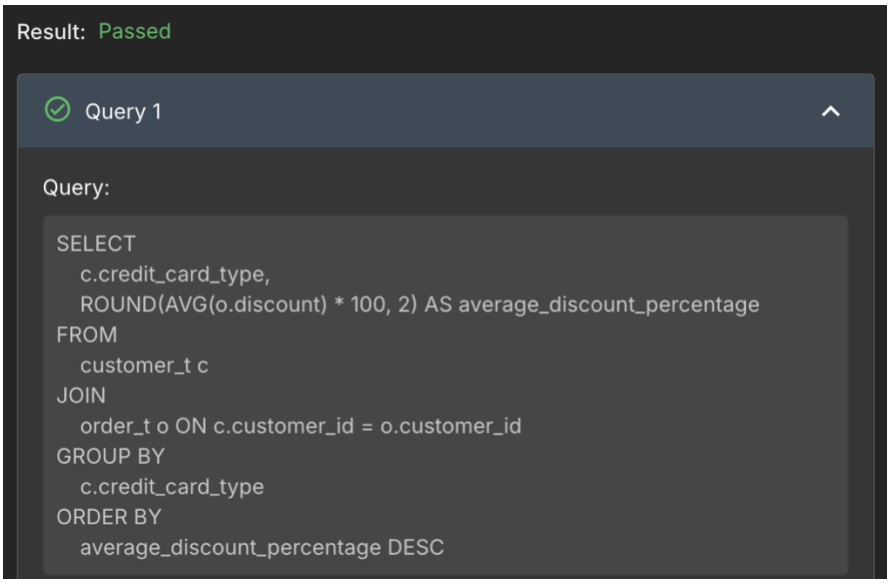
- Q1 had the highest net revenue and total of orders indicating a strong start to the year.
- Q3 sees the most significant revenue drop (~32.3%) from Q2.
- Over the course of the year, total revenue drops by over 52%.

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

Solution Query:

```
SELECT
    c.credit_card_type,
    ROUND(AVG(o.discount) * 100, 2) AS average_discount_percentage
FROM
    customer_t c
JOIN
    order_t o ON c.customer_id = o.customer_id
GROUP BY
    c.credit_card_type
ORDER BY
    average_discount_percentage DESC;
```

Output:



Output:

Showing first 10 rows out of 16 rows

credit_card_type	average_discount_per...
laser	64.38
mastercard	62.95
maestro	62.42
visa-electron	62.35
china-unionpay	62.22
instapayment	62.06
americanexpress	61.63
diners-club-us-ca	61.46
diners-club-carte-blan...	61.45
switch	61.02

Observations and Insights:

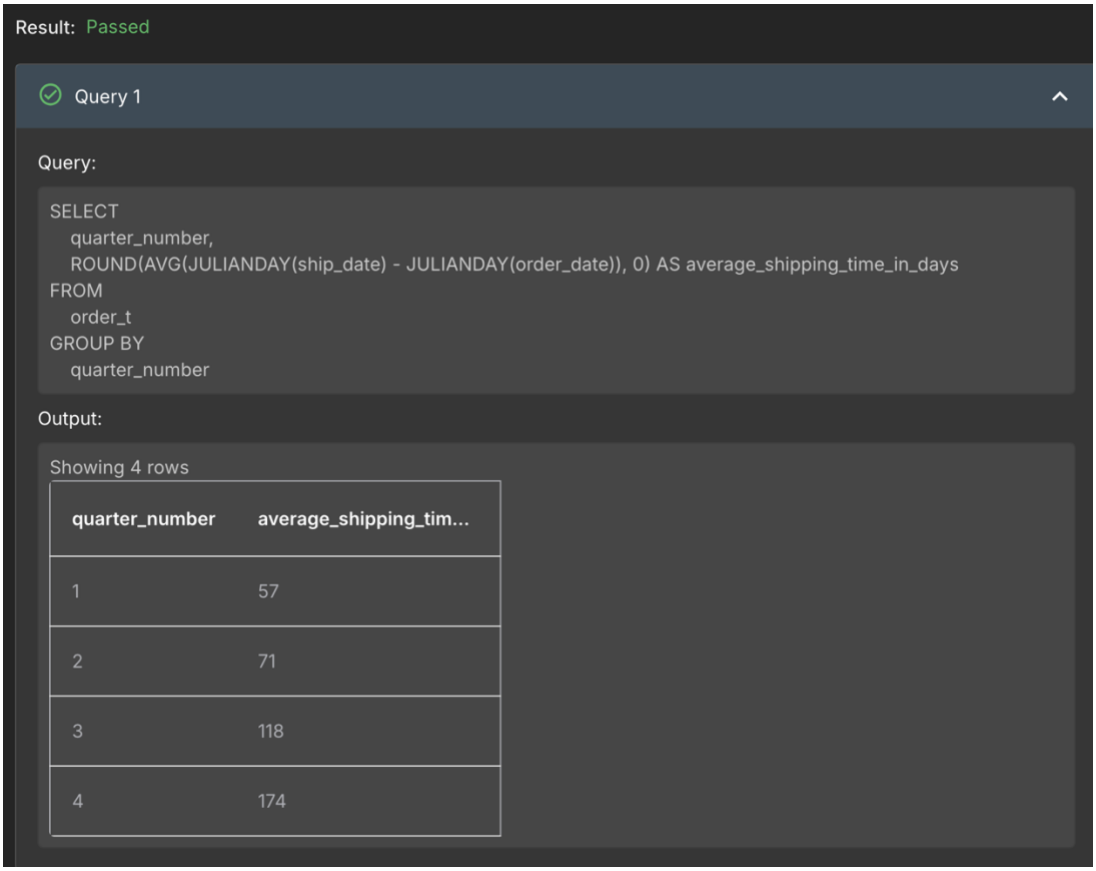
- The top 10 credit card types have average discounts ranging between 61.02%-64.38% reflecting small variations in discount percentages between card types.
- The Laser credit card had the highest discount indicating that customers using this card type are saving more.
- Both Diners Club cards have extremely similar average discounts.

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

Solution Query:

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

Output:



Observations and Insights:

- Average shipping time increases over time suggesting potential inefficiencies or delays in logistics.
- The shipping time in Q4 is approximately 3 times longer than in Q1.

- The increase in shipping time, particularly in Q3 and Q4, could be the reason for rising customer dissatisfaction.

Business Metrics Overview

Total Revenue	Total Orders	Total Customers	Average Rating
$ 48,610,993.78	1000	994	Okay
Last Quarter Revenue	**Last quarter Orders**	**Average Days to Ship**	**% Good Feedback**
$8,573,149.28	199	98	21.5%

Business Recommendations

1. There has been a notable rise in shipping times, rising from 57 days in Q1 to 174 days in Q4, which directly correlated with increased customer dissatisfaction.
   Recommendations:
   - Prioritize enhancing logistics by partnering with dependable shipping providers to streamline operations.
   - Display estimated delivery times on the app to set realistic expectations for customers and build trust.
   - Regularly monitor key shipping metrics and aim to reduce shipping times within the next 2 quarters.
2. Orders have decreased quarter over quarter, with the most significant decline of 15.5% occurring between Q1 and Q2.
   Recommendations:
   - Create seasonal promotions or incentives during low-performing quarters.
   - Partner with third-party financing companies to offer more flexible payment options.
   - Encourage repeat business by providing exclusive trade-in bonuses for returning customers.
3. Over half of the customers in Q4 reported dissatisfaction, with a significant (~47%) increase in dissatisfaction from Q3 to Q4.
   Recommendations
   - Conduct customer satisfaction surveys/follow-ups after vehicle deliveries to gain insight on pain points and address any issues promptly.
   - Establish quarterly targets to boost customer satisfaction ratings by at least 10%.
   - Encourage satisfied customers to leave positive reviews by offering small perks for participating in feedback initiatives.