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.

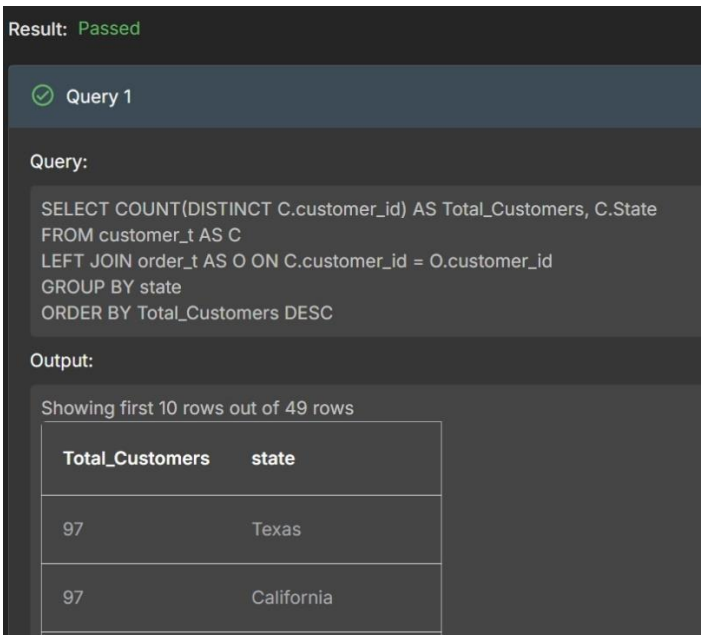
Business Questions

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

Solution Query:

```
SELECT COUNT(DISTINCT C.customer_id) AS Total_Customers, C.State

FROM customer_t AS C

LEFT JOIN order_t AS O ON C.customer_id = O.customer_id

GROUP BY state

ORDER BY Total_Customers DESC;
```

Output:



97	California
86	Florida
69	New York
35	District of Columbia
33	Ohio
33	Colorado
29	Alabama
28	Washington
26	Arizona

Observations and Insights:

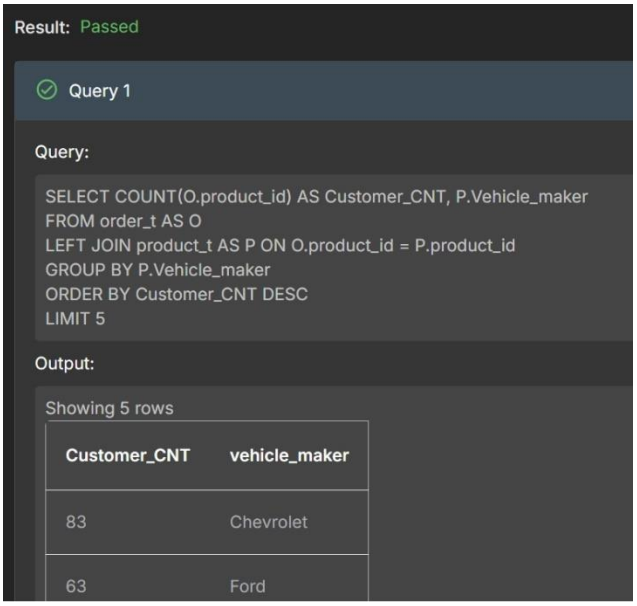
- Texas, California, Florida have the most customers who've placed orders which may mean a higher population in these areas, better access to these kinds of services, or a better market size.
- Businesses might focus on states like Texas and California for marketing or expanding services due to their large customer base.
- States with lower customer counts may represent untapped or underperforming markets.

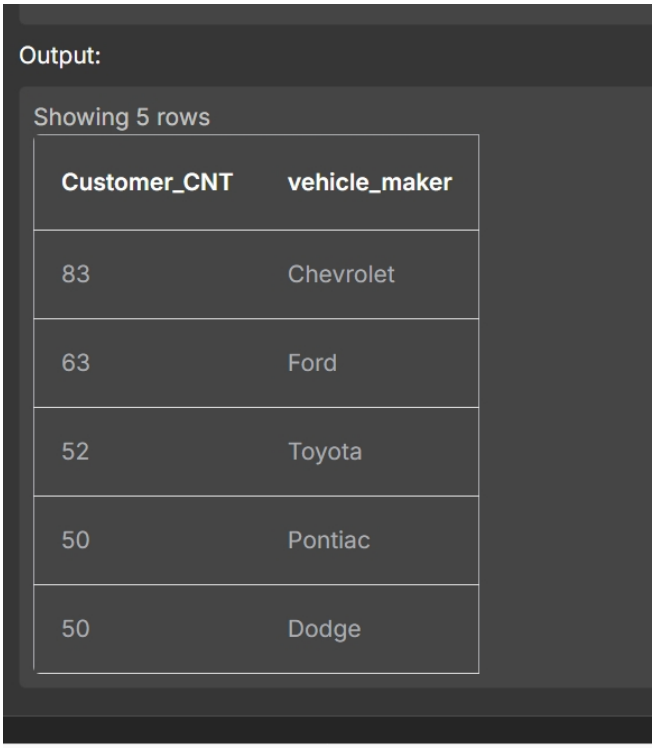
Question 2: Which are the top 5 vehicle makers preferred by the customers?

Solution Query:

```
SELECT COUNT(O.product_id) AS Customer_CNT, P.Vehicle_maker

FROM order_t AS O

LEFT JOIN product_t AS P ON O.product_id = P.product_id

GROUP BY P.Vehicle_maker

ORDER BY Customer_CNT DESC

LIMIT 5;
```

Output:





Observations and Insights:

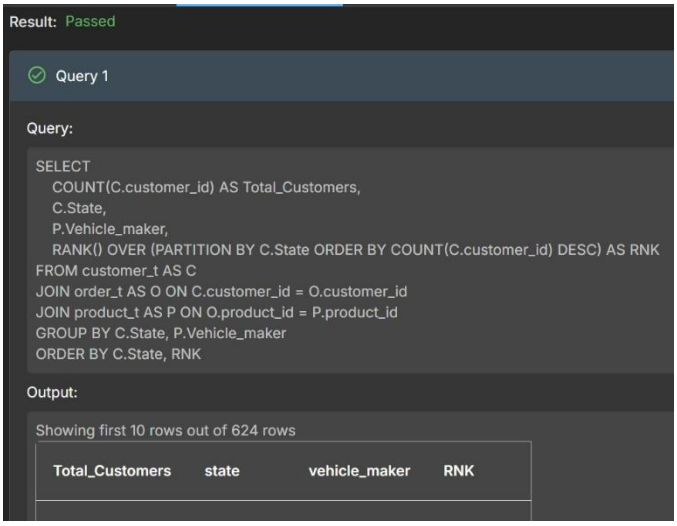
- Preference for vehicle brands (Chevrolet, Ford, and Dodge) over international brands like Toyota, though Toyota maintains a strong position in third place.
- Chevrolet: Leads with 83 customers., Ford: Comes second with 63 customers, Toyota: Ranks third with 52 customers.
- Chevrolet dominates as the most preferred brand, indicating strong brand loyalty or appeal.

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

Solution Query:

```
SELECT
    COUNT(C.customer_id) AS Total_Customers,
    C.State,
    P.Vehicle_maker,
    RANK() OVER (PARTITION BY C.State ORDER BY COUNT(C.customer_id) DESC) AS RNK
FROM customer_t AS C
JOIN order_t AS O ON C.customer_id = O.customer_id
JOIN product_t AS P ON O.product_id = P.product_id
GROUP BY C.State, P.Vehicle_maker
ORDER BY C.State, RNK;
```

Output:



Total_Customers	state	vehicle_maker	RNK
5	Alabama	Dodge	1
3	Alabama	Ford	2
2	Alabama	Plymouth	3
2	Alabama	Mazda	3
2	Alabama	Chevrolet	3
1	Alabama	Volvo	6
1	Alabama	Volkswagen	6
1	Alabama	Suzuki	6

1	Alabama	Volvo	6
1	Alabama	Volkswagen	6
1	Alabama	Suzuki	6
1	Alabama	Subaru	6
1	Alabama	Porsche	6

Observations and Insights:

- In Alabama, Dodge appears to be the most preferred vehicle maker, with a rank of 1 and a customer count of 5.
- There is a tie among Plymouth, Mazda, and Chevrolet for rank 3, each with 2 customers.
- Alabama shows a wide variety of vehicle makers with ranks extending to 6, indicating a diverse preference base.

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 ROUND(AVG(feedback), 2) AS average_feedback, quarter_number
FROM (
   SELECT
      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 feedback,
      quarter_number
   FROM order_t
) AS feedback_data
GROUP BY quarter_number;
```

Output:

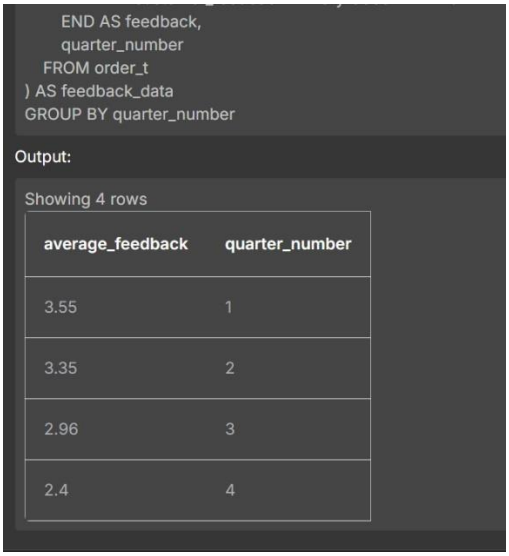
Result: Passed

Query 1

Query:

```
SELECT ROUND(AVG(feedback), 2) AS average_feedback, quarter_number
FROM (
  SELECT
    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 feedback,
    quarter_number
  FROM order_t
) AS feedback_data
GROUP BY quarter_number
```

Output:



Observations and Insights:

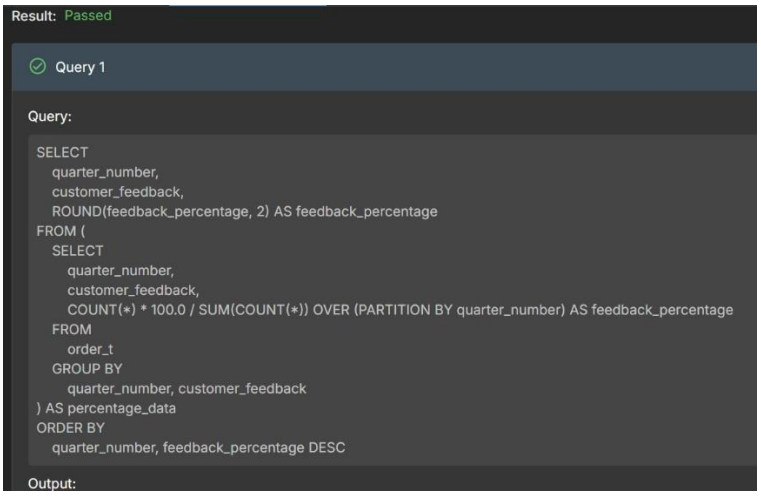
- The average feedback score decreases steadily across the quarters, this suggests a downward trend in customer satisfaction or experience over the course of the year
- The most noticeable decline occurs between Quarter 3 (2.96) and Quarter 4 (2.40), which indicates a sharp reduction in feedback scores during this period.
- Efforts should be made to investigate what caused the significant decline in feedback in the quarters.

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

Solution Query:

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

Output:



Output:

Showing first 10 rows out of 20 rows

quarter_number	customer_feedback	feedback_percentage
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

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

Observations and Insights:

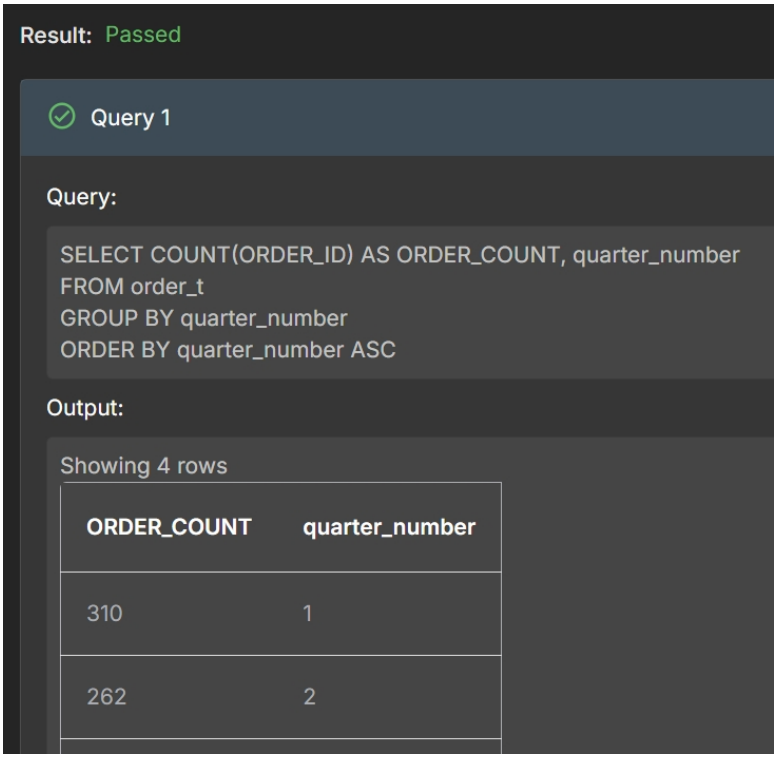
- At the beginning of each quarter the feedback percentage decreases, for example quarter one goes from 30% to 10.97% and quarter two goes from 28.63% to 14.12%. Meaning customers are getting more dissatisfied as the quarter progresses.
- The increase in dissatisfaction suggests possible issues in the customer experience.
- Are there specific areas (products, services, teams) contributing to the negative feedback?

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

Solution Query:

```
SELECT COUNT(ORDER_ID) AS ORDER_COUNT, quarter_number

FROM order_t

GROUP BY quarter_number

ORDER BY quarter_number ASC;
```

Output:



```
SELECT COUNT(ORDER_ID) AS ORDER_COUNT, quarter_number
FROM order_t
GROUP BY quarter_number
ORDER BY quarter_number ASC
```

Output:

Showing 4 rows

ORDER_COUNT	quarter_number
310	1
262	2
229	3
199	4

Observations and Insights:

- There is a decline in orders as the year progresses.
- Quarter 1 has the highest number of orders (310). This could indicate a strong start to the year, likely due to factors like New Year sales, post-holiday demand, or effective marketing campaigns.
- decline could reflect seasonal factors, waning customer interest, or external challenges like economic slowdowns, reduced inventory, or lower advertising.

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,
    net_revenue,
    previous_revenue,
    ROUND(((net_revenue - previous_revenue) / NULLIF(previous_revenue, 0)) * 100, 2) AS qoq_percentage_change
FROM (
    SELECT
        quarter_number,
        SUM(quantity * (COALESCE(vehicle_price, 0) - COALESCE(discount, 0))) AS net_revenue,
        LAG(SUM(quantity * (COALESCE(vehicle_price, 0) - COALESCE(discount, 0))))
        OVER (ORDER BY quarter_number) AS previous_revenue
    FROM
        order_t
    GROUP BY
        quarter_number
) AS revenue_with_lag
ORDER BY
    quarter_number;
```

Output:

Result: Passed

Query 1

Query:

```
SELECT
    quarter_number,
    net_revenue,
    previous_revenue,
    ROUND(((net_revenue - previous_revenue) / NULLIF(previous_revenue, 0)) * 100, 2) AS qoq_
FROM (
    SELECT
        quarter_number,
        SUM(quantity * (COALESCE(vehicle_price, 0) - COALESCE(discount, 0))) AS net_revenue,
        LAG(SUM(quantity * (COALESCE(vehicle_price, 0) - COALESCE(discount, 0))))
        OVER (ORDER BY quarter_number) AS previous_revenue
    FROM
        order_t
    GROUP BY
        quarter_number
) AS revenue_with_lag
ORDER BY
    quarter_number
LIMIT 5
```

Output:

Showing 4 rows

quarter_number	net_revenue	previous_revenue	qoq_percentage_change
1	39637378.160000026		
2	32913497.43999998	39637378.160000026	-16.96
3	29435188.489999995	32913497.43999998	-10.57
4	23495814.140000004	29435188.489999995	-20.18

Observations and Insights:

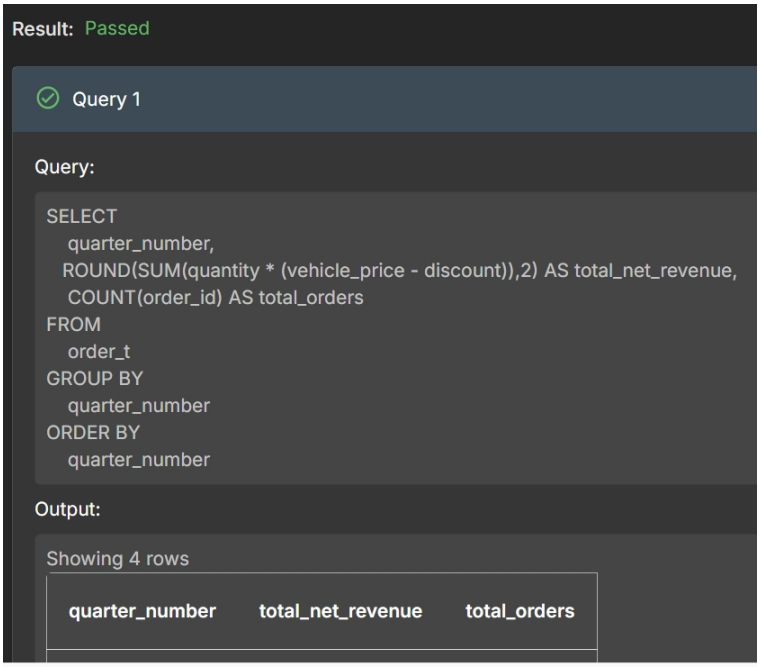
- Net revenue shows a steady decline over all quarters, this could indicate persistent issues such as reduced demand, operational inefficiencies, or increased competition.
- The sharp drop in Quarter 4 (20.18%) suggests worsening performance. Possible causes include seasonality, higher discounts affecting revenue, or loss of customers.
- Address the decline by identifying high-performing products or services and scaling efforts around them. Revise discount strategies to balance competitiveness and profitability.

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

Solution Query:

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

Output:



Output:

Showing 4 rows

quarter_number	total_net_revenue	total_orders
1	39637378.16	310
2	32913497.44	262
3	29435188.49	229
4	23495814.14	199

Observations and Insights:

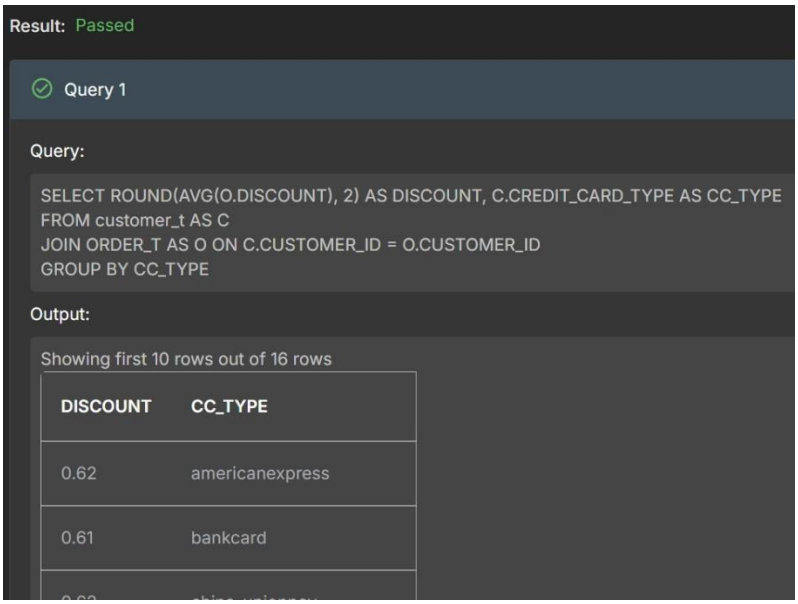
- There's a decline in total net revenue and total order as the quarters progresses.
- Both total net revenue and total orders decline quarter over quarter, indicating fewer transactions and sales activity overall.
- The consistent decrease in total orders suggests a declining demand for products or services over the year.

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

Solution Query:

```
SELECT ROUND(AVG(O.DISCOUNT), 2) AS DISCOUNT, C.CREDIT_CARD_TYPE AS CC_TYPE
FROM customer_t AS C
JOIN ORDER_T AS O ON C.CUSTOMER_ID = O.CUSTOMER_ID
GROUP BY CC_TYPE;
```

Output:



0.61	bankcard
0.62	china-unionpay
0.61	diners-club-carte-blan...
0.6	diners-club-enroute
0.58	diners-club-international
0.61	diners-club-us-ca
0.62	instapayment
0.61	jcb
0.64	laser

Observations and Insights:

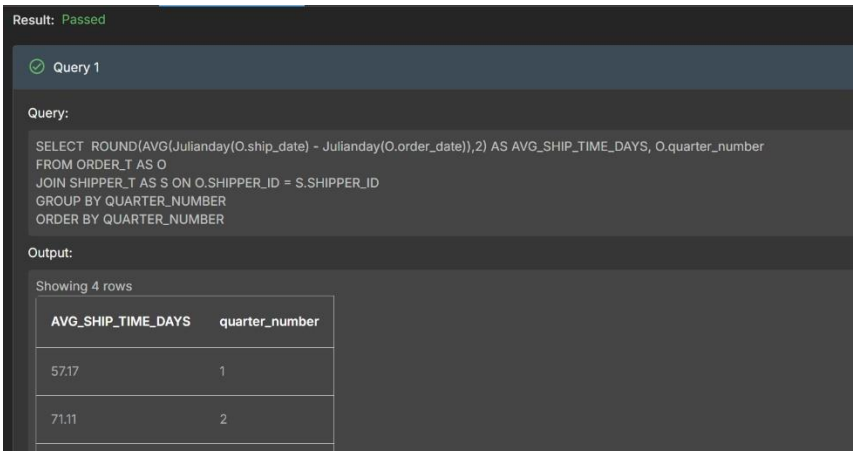
- Most credit card types have an average discount close to 0.60–0.62, indicating a relatively consistent discount strategy for most customers.
- The average discount varies slightly across credit card types, ranging from 0.58 to 0.64., The highest discount is offered for customers using the Laser card (0.64)., The lowest discount is for Diners Club International (0.58).
- Laser card users may receive preferential discounts, which could be due to partnerships or promotional strategies. Diners Club International users might receive lower average discounts, potentially reflecting lower usage or fewer promotional incentives.

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

Solution Query:

```
SELECT  ROUND(AVG(Julianday(O.ship_date) - Julianday(O.order_date)),2) AS AVG_SHIP_TIME_DAYS,
O.quarter_number

FROM ORDER_T AS O

JOIN SHIPPER_T AS S ON O.SHIPPER_ID = S.SHIPPER_ID

GROUP BY QUARTER_NUMBER

ORDER BY QUARTER_NUMBER;
```

Output:



```
JOIN SHIPPER_T AS S ON O.SHIPPER_ID = S.SHIPPER_ID
GROUP BY QUARTER_NUMBER
ORDER BY QUARTER_NUMBER
```

Output:

Showing 4 rows

AVG_SHIP_TIME_DAYS	quarter_number
57.17	1
71.11	2
117.76	3
174.1	4

Observations and Insights:

- The shipping time more than triples from Q1 to Q4, indicating a significant delay as the year progresses.
- Higher demand in later quarters like the holiday seasons may strain resources, leading to longer shipping times.
- . Longer shipping times in Q3 and Q4 could negatively impact customer satisfaction, especially if delivery expectations are not met.

Business Metrics Overview

Total Revenue	Total Orders	Total Customers	Average Rating
$126,382,378.22	1,000	994	2.95
Last Quarter Revenue	**Last quarter Orders**	**Average Days to Ship**	**% Good Feedback**
$23,495,814.14	**199**	97.96	44.1

Business Recommendations

- Address the decline in customer feedback scores, particularly the drop in "Good" and "Very Good" feedback over time. Conduct surveys or gather insights to identify the root causes of dissatisfaction
- States like Texas, California, and Florida have a larger customer base. Focus marketing efforts in these regions with targeted ads and promotions.
- States with fewer customers represent untapped or underperforming markets. Analyze customer behavior and preferences in these regions and offer tailored promotions, discounts, or partnerships to attract new customers.
- Brands like Chevrolet, Ford, and Toyota dominate customer preferences. Promote these brands more prominently on the platform.
- Both revenue and orders decline quarter over quarter. Look into reasons like seasonal changes, losing customers, or competition. Run special promotions during slow quarters to increase sales. Use data to find customers who might leave and offer them incentives to stay.