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.

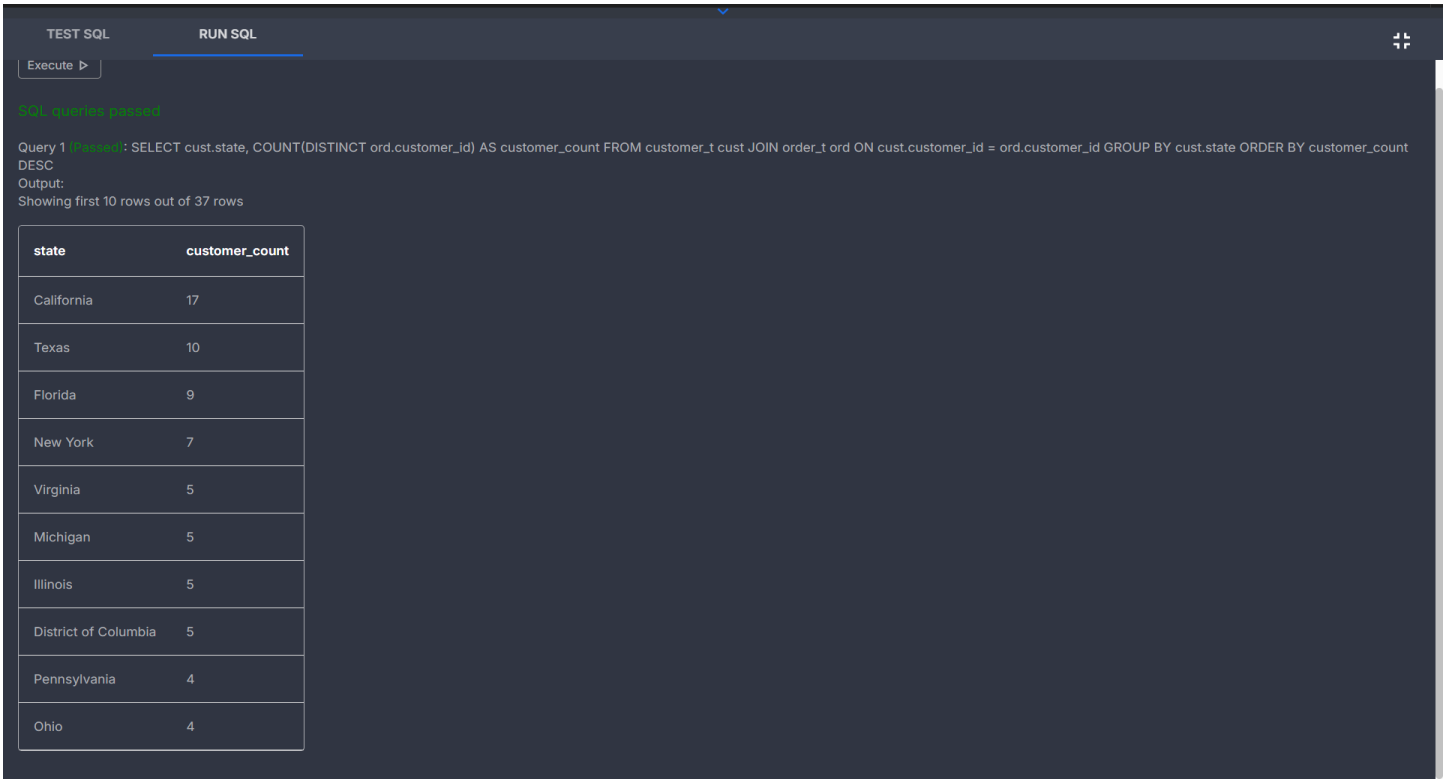
Business Questions

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

Solution Query:

```
SELECT
        cust.state,
        COUNT(DISTINCT ord.customer_id) AS customer_count
FROM
        customer_t cust
JOIN
        order_t ord ON cust.customer_id = ord.customer_id
GROUP BY
        cust.state
ORDER BY
        customer_count DESC;
```

Output:



state	customer_count
California	17
Texas	10
Florida	9
New York	7
Virginia	5
Michigan	5
Illinois	5
District of Columbia	5
Pennsylvania	4
Ohio	4

Observations and Insights:

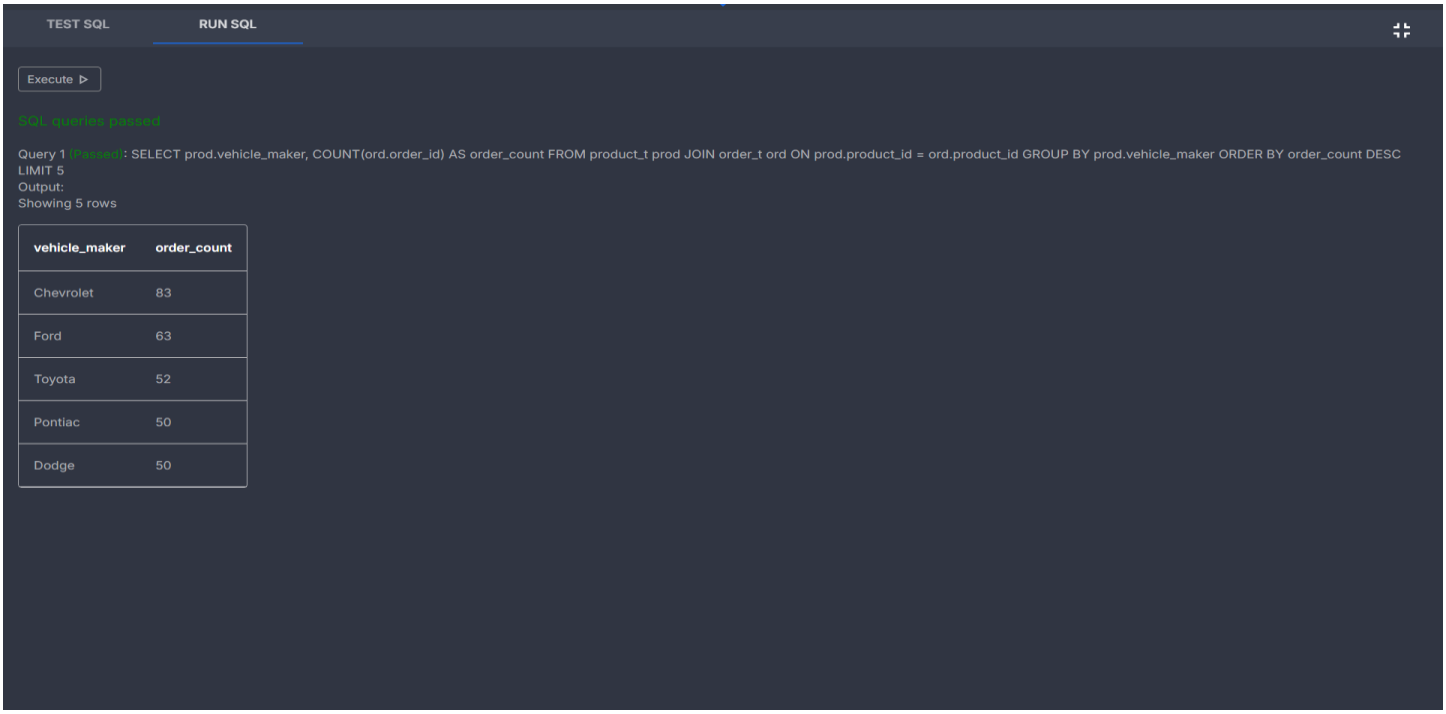
- California has the highest number of customers, with 17 placing orders.
- Texas and Florida follow with 10 and 9 customers, respectively.
- New York ranks fourth, with 7 customers placing orders.
- Virginia, Michigan, Illinois, and the District of Columbia each have 5 customers.
- Pennsylvania has fewer customers, with only 4 placing orders.
- California shows strong market presence, indicating potential for continued focus.
- Texas and Florida represent important markets for growth and customer engagement.
- New York presents an opportunity for expansion with targeted marketing efforts.
- Virginia, Michigan, Illinois, and D.C. are mid-tier markets with potential for increased focus.
- Pennsylvania's lower customer count suggests room for market expansion.

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

Solution Query:

```
SELECT
        prod.vehicle_maker,
        COUNT(ord.order_id) AS order_count
FROM
        product_t prod
JOIN
        order_t ord ON prod.product_id = ord.product_id
GROUP BY
        prod.vehicle_maker
ORDER BY
        order_count DESC
LIMIT 5;
```

Output:



Observations and Insights:

- Chevrolet's dominance suggests that New-Wheels' customer base has a strong preference for this brand, possibly due to its reputation, pricing, or availability of models.
- Ford and Toyota also hold significant market share, indicating that focusing on these brands could drive further sales.
- Pontiac and Dodge being tied for fourth place shows customer interest in mid-range brands, and New-Wheels might benefit from promoting these brands more effectively.
- Targeted marketing for Chevrolet, Ford, and Toyota vehicles could further boost sales, while exploring promotional strategies for Pontiac and Dodge could help maintain their steady customer demand.

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

Solution Query:

```
SELECT
        state,
        vehicle_maker,
        order_count
FROM (
        SELECT
                cust.state,
                prod.vehicle_maker,
                COUNT(ord.order_id) AS order_count,
                RANK() OVER (PARTITION BY cust.state ORDER BY COUNT(ord.order_id) DESC) AS rank
        FROM
                customer_t cust
        JOIN
                order_t ord ON cust.customer_id = ord.customer_id
        JOIN
                product_t prod ON ord.product_id = prod.product_id
        GROUP BY
                cust.state, prod.vehicle_maker
        ) AS ranked_vehicle_makers
WHERE
        rank = 1
ORDER BY
        order_count desc;
```

Output:

TEST SQL RUN SQL

SQL queries passed

Query 1 (Passed): SELECT state, vehicle_maker, order_count FROM (SELECT cust.state, prod.vehicle_maker, COUNT(ord.order_id) AS order_count, RANK() OVER (PARTITION BY cust.state ORDER BY COUNT(ord.order_id) DESC) AS rank FROM customer_t cust JOIN order_t ord ON cust.customer_id = ord.customer_id JOIN product_t prod ON ord.product_id = prod.product_id GROUP BY cust.state, prod.vehicle_maker) AS ranked_vehicle_makers WHERE rank = 1 ORDER BY order_count desc
Output:
Showing first 10 rows out of 101 rows

state	vehicle_maker	order_count
California	Pontiac	2
California	Nissan	2
California	Ford	2
California	Chevrolet	2
Florida	Volvo	2
Florida	Ford	2
Indiana	Mazda	2
Texas	Nissan	2
Alabama	Lincoln	1
Alabama	Lexus	1

Observations and Insights:

- California shows a highly competitive vehicle market with Pontiac, Nissan, Ford, and Chevrolet all having an equal order count of 2, indicating no clear dominant vehicle maker.
- Florida has a split preference between Volvo and Ford, both receiving 2 orders, highlighting a balanced demand for these brands.
- Indiana sees Mazda as the most preferred vehicle maker, with 2 orders, showing a localized preference for this brand.
- In Texas, Nissan leads with 2 orders, suggesting a modest but significant presence in the state's vehicle market.
- Alabama shows a more fragmented preference with Lincoln and Lexus each having only 1 order, indicating lower market activity and a dispersed customer preference.
- The results suggest that multiple vehicle makers are tied with the same number of orders in various states, reflecting diverse consumer preferences across regions.
- The low order counts across most states, even for the top vehicle makers, may signal a need for greater market penetration or marketing efforts to increase engagement.
- States with equal counts for different vehicle makers suggest an opportunity for competitive strategies to capture market share in these regions.

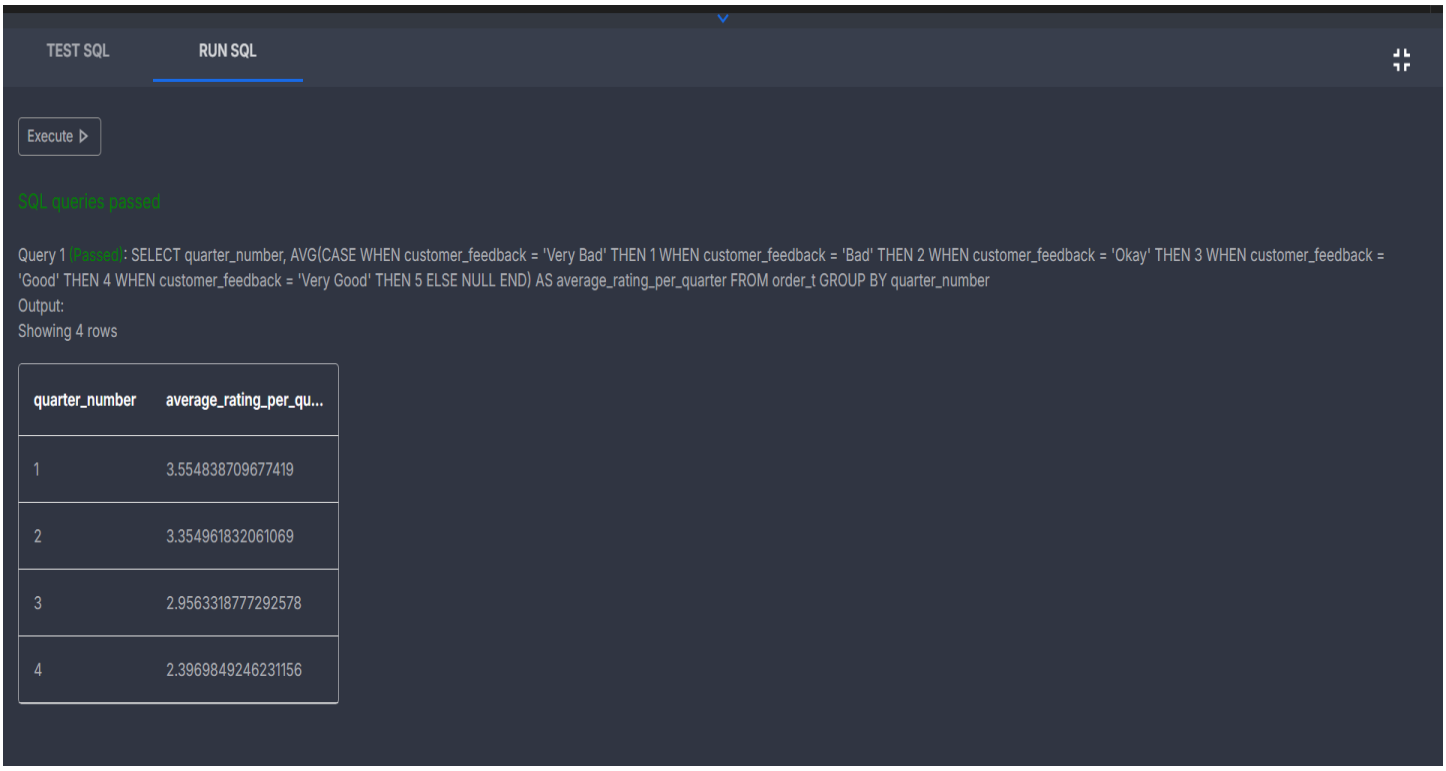
Question 4: Find the overall average rating given by the customers. What is the average rating in each quarter?

Consider the following mapping for ratings: “Very Bad”: 1, “Bad”: 2, “Okay”: 3, “Good”: 4, “Very Good”: 5

Solution Query:

```
SELECT
   quarter_number,
   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
      ELSE NULL
   END) AS average_rating_per_quarter
FROM order_t
GROUP BY quarter_number;
```

Output:



Observations and Insights:

- Quarter 1 has the highest average rating of 3.55, indicating that customers were generally more satisfied with their purchases and services during this period compared to other quarters. The rating falls between "Okay" and "Good."
- Quarter 2 shows a slight decline in average rating to 3.35, still within the "Okay" to "Good" range, but lower than in Quarter 1, indicating a small drop in customer satisfaction.
- Quarter 3 continues the downward trend, with an average rating of 2.96, indicating a shift closer to the "Okay" rating and suggesting that customer satisfaction is diminishing significantly as the year progresses.
- Quarter 4 has the lowest average rating of 2.40, meaning customer satisfaction drops to between "Bad" and "Okay." This indicates a serious decline in customer experiences and could be a critical point for the business.
- The continuous decline in customer ratings from Quarter 1 to Quarter 4 shows a worsening trend in customer satisfaction. This downward trajectory may contribute to the company's overall drop in sales and new customer acquisition. Immediate action to improve after-sales service and product quality is crucial, especially as ratings in the last quarter approach a "Bad" level.

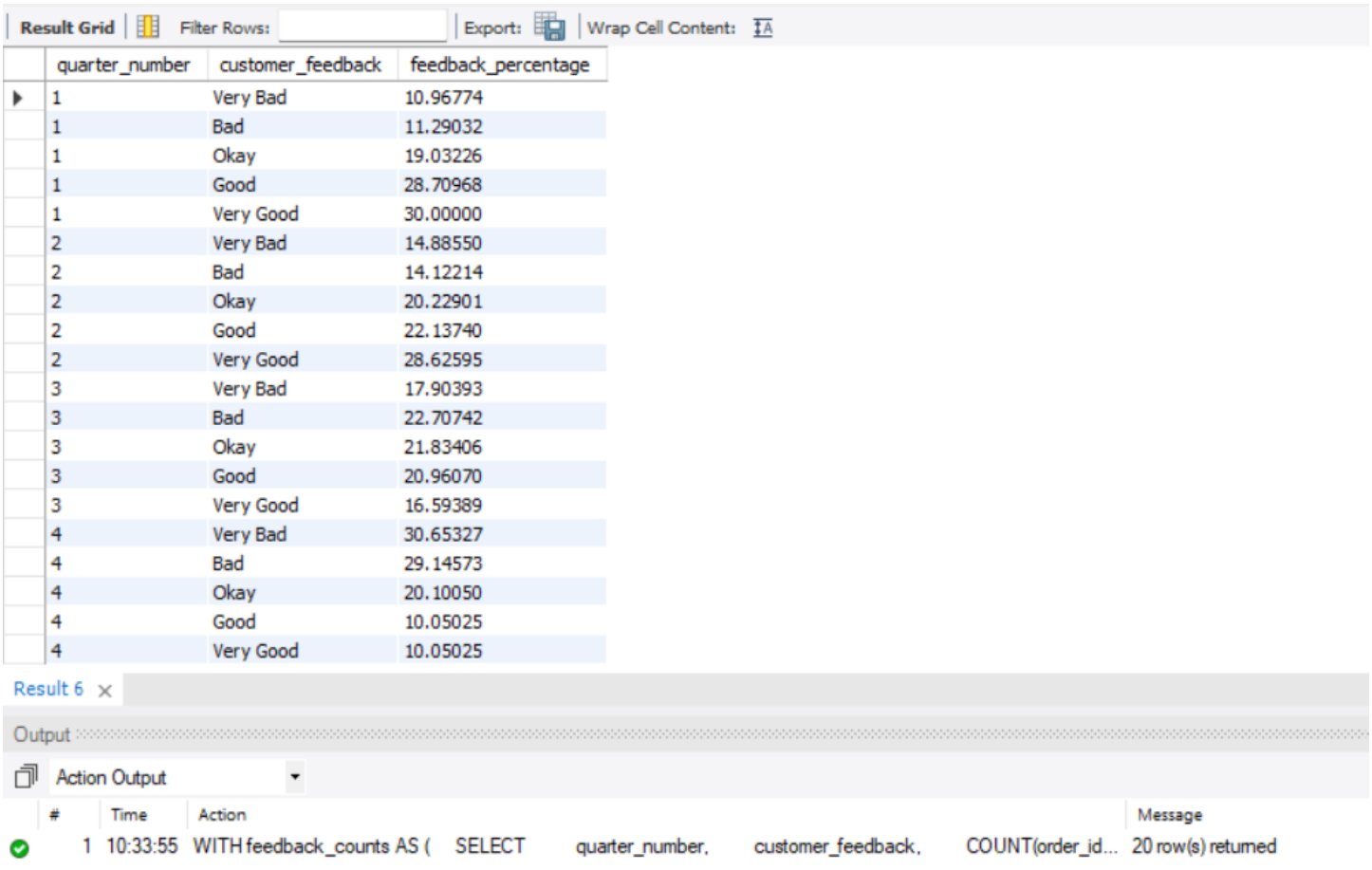
Question 5: Find the percentage distribution of feedback from the customers. Are customers getting more dissatisfied over time?

Solution Query:

```
WITH feedback_counts AS (
   SELECT
      quarter_number,
      customer_feedback,
      COUNT(order_id) AS feedback_count,
      COUNT(order_id) * 100.0 / SUM(COUNT(order_id)) OVER (PARTITION BY quarter_number) AS
feedback_percentage
   FROM
      order_t
   GROUP BY
      quarter_number, customer_feedback
)

SELECT
   quarter_number,
   customer_feedback,
   feedback_percentage
FROM
   feedback_counts
ORDER BY
   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;
```

Output:



Result Grid | Filter Rows: | Export: | Wrap Cell Content:

quarter_number	customer_feedback	feedback_percentage
1	Very Bad	10.96774
1	Bad	11.29032
1	Okay	19.03226
1	Good	28.70968
1	Very Good	30.00000
2	Very Bad	14.88550
2	Bad	14.12214
2	Okay	20.22901
2	Good	22.13740
2	Very Good	28.62595
3	Very Bad	17.90393
3	Bad	22.70742
3	Okay	21.83406
3	Good	20.96070
3	Very Good	16.59389
4	Very Bad	30.65327
4	Bad	29.14573
4	Okay	20.10050
4	Good	10.05025
4	Very Good	10.05025

Result 6

Output

Action Output

#	Time	Action	Message
1	10:33:55	WITH feedback_counts AS (SELECT quarter_number, customer_feedback, COUNT(order_id...	20 row(s) returned

Observations and Insights:

1. Quarter 1 Feedback Distribution:
   - Majority of customers (58.71%) gave positive feedback ('Good' and 'Very Good').
   - Negative feedback ('Very Bad' and 'Bad') combined was 22.26%, indicating a relatively balanced mix of satisfaction levels.
2. Quarter 2 Feedback Distribution:
   - There was a slight increase in negative feedback ('Very Bad' and 'Bad'), which reached 29%, compared to 22.26% in Quarter 1.
   - Positive feedback ('Good' and 'Very Good') dropped slightly to 50.76%.
3. Quarter 3 Feedback Distribution:
   - A significant rise in negative feedback, with 40.61% of feedback being 'Very Bad' or 'Bad.'
   - Positive feedback dropped to 37.55%, reflecting an increase in customer dissatisfaction.
4. Quarter 4 Feedback Distribution:
   - A sharp increase in dissatisfaction, with negative feedback ('Very Bad' and 'Bad') at 59.8%, up from 40.61% in Quarter 3.
   - Positive feedback ('Good' and 'Very Good') drastically declined to just 20.1%, highlighting major concerns in customer experience.

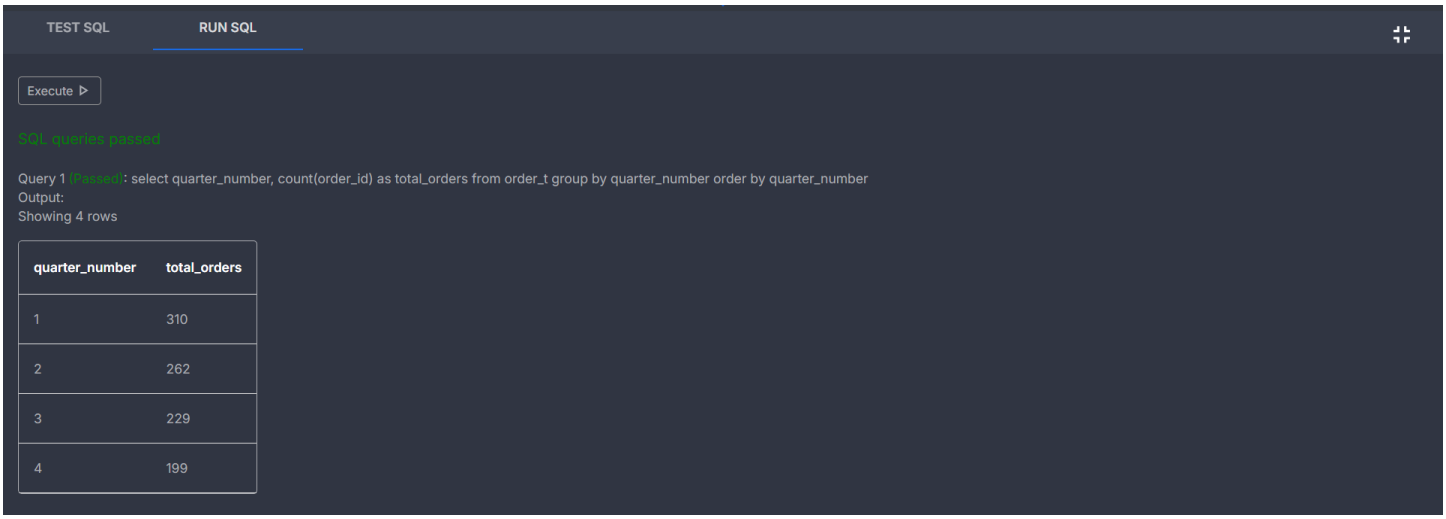
- Customer dissatisfaction has been steadily increasing, with negative feedback sharply rising from Quarter 1 (22.26%) to Quarter 4 (59.8%).
- Positive feedback significantly declined, particularly in the last quarter, indicating potential problems with product quality or service during this period.
- A deeper investigation is required into why customer satisfaction has worsened over time, especially in Quarter 4.

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

Solution Query:

```
select
   quarter_number,
   count(order_id) as total_orders
from
   order_t
group by
   quarter_number
order by
   quarter_number;
```

Output:



Observations and Insights:

- The number of total orders decreases each quarter over the year.
- Quarter 1 has the highest number of orders (310), showing strong performance at the start of the year.
- Quarter 4 has the lowest number of orders (199), indicating a significant decline compared to Quarter 1.
- The trend shows a steady decline in customer activity or order volume as the year progresses.
- The drop in orders from Quarter 2 to Quarter 3 is less pronounced compared to the drop from Quarter 1 to Quarter 2.

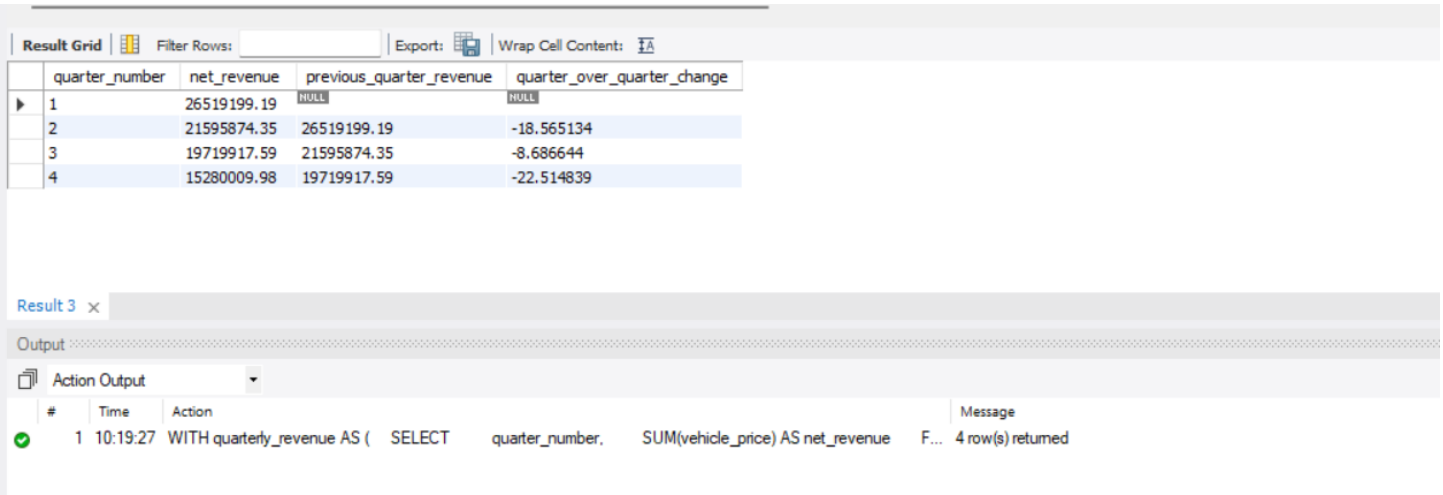
Question 7: Calculate the net revenue generated by the company. What is the quarter-over-quarter % change in net revenue?

Solution Query:

```
WITH quarterly_revenue AS (
   SELECT
      quarter_number,
      SUM(vehicle_price) AS net_revenue
   FROM
      order_t  -- Using order_t directly for revenue
   GROUP BY
      quarter_number
)

SELECT
   q.quarter_number,
   q.net_revenue,
   LAG(q.net_revenue) OVER (ORDER BY q.quarter_number) AS previous_quarter_revenue,
   CASE
      WHEN LAG(q.net_revenue) OVER (ORDER BY q.quarter_number) IS NOT NULL
      THEN ((q.net_revenue - LAG(q.net_revenue) OVER (ORDER BY q.quarter_number)) /
LAG(q.net_revenue) OVER (ORDER BY q.quarter_number)) * 100
      ELSE NULL
   END AS quarter_over_quarter_change
FROM
   quarterly_revenue q
ORDER BY
   q.quarter_number;
```

Output:



Result Grid | Filter Rows: | Export: | Wrap Cell Content:

quarter_number	net_revenue	previous_quarter_revenue	quarter_over_quarter_change
1	26519199.19	NULL	NULL
2	21595874.35	26519199.19	-18.565134
3	19719917.59	21595874.35	-8.686644
4	15280009.98	19719917.59	-22.514839

Result 3

Output

Action Output

#	Time	Action	Message
1	10:19:27	WITH quarterly_revenue AS (SELECT quarter_number, SUM(vehicle_price) AS net_revenue F...	4 row(s) returned

Observations and Insights:

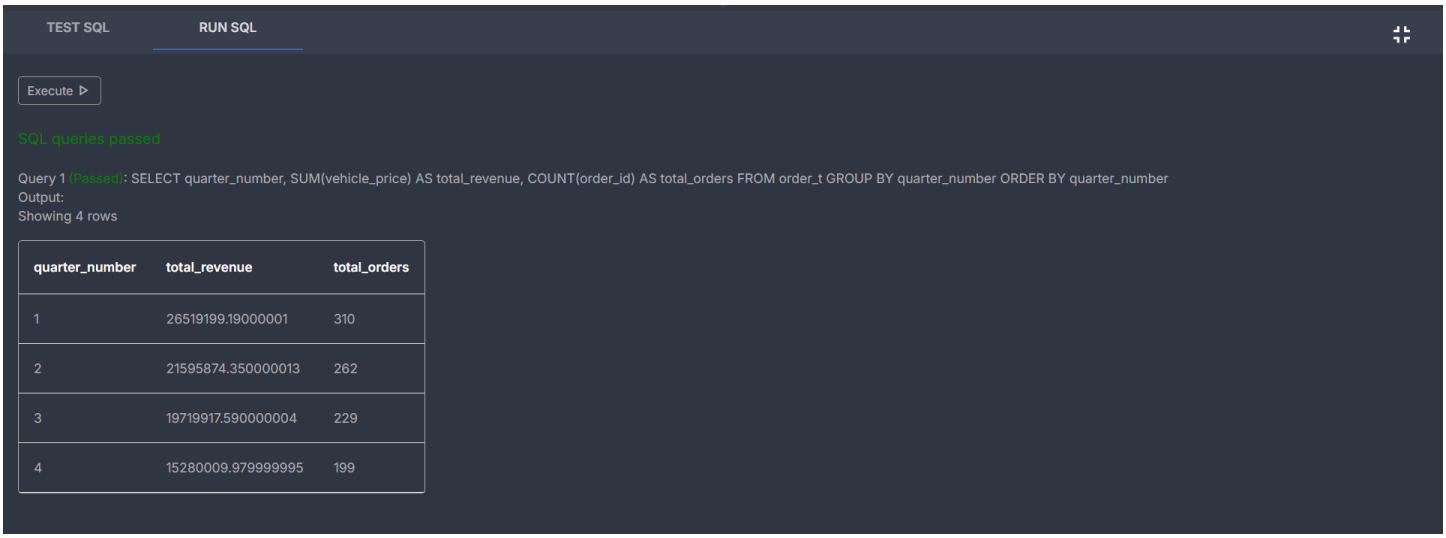
- Q1 Net Revenue: $26,519,199.19 - This represents the highest quarterly revenue but serves as the baseline with no prior comparison.
- Q2 Revenue Decline: Net revenue dropped by 18.57% compared to Q1, indicating a significant decrease in sales during the second quarter.
- Q3 Slower Decline: Revenue fell by 8.69% from Q2 to Q3, suggesting a slower decline compared to the previous quarter, but still showing downward pressure.
- Q4 Significant Revenue Drop: The largest drop occurred in Q4 with a 22.51% decrease in revenue compared to Q3, indicating a sharp decline in sales performance during this period.
- Overall Trend: The consistent revenue decline across all quarters points to potential issues with maintaining customer demand or market conditions, requiring deeper investigation into sales strategies, market trends, and customer engagement efforts.

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

Solution Query:

```
SELECT
   quarter_number,
   SUM(vehicle_price) AS total_revenue,
   COUNT(order_id) AS total_orders
FROM
   order_t
GROUP BY
   quarter_number
ORDER BY
   quarter_number;
```

Output:



Observations and Insights:

- Decreasing Trend in Total Revenue: There is a noticeable decline in total revenue from Quarter 1 to Quarter 4, dropping from approximately $26.5 million in Q1 to about $15.3 million in Q4. This indicates a potential downward trend in sales over the year.
- Declining Number of Orders: The total number of orders also decreased from Q1 to Q4, starting at 310 orders in Q1 and falling to 199 orders in Q4. This trend suggests that fewer customers are making purchases as the year progresses.
- Quarterly Revenue Drop: The largest decrease in revenue occurs from Q1 to Q2, where the revenue drops by about 18.6%. Subsequent quarters also show decreases, with the largest drop occurring in Q4.

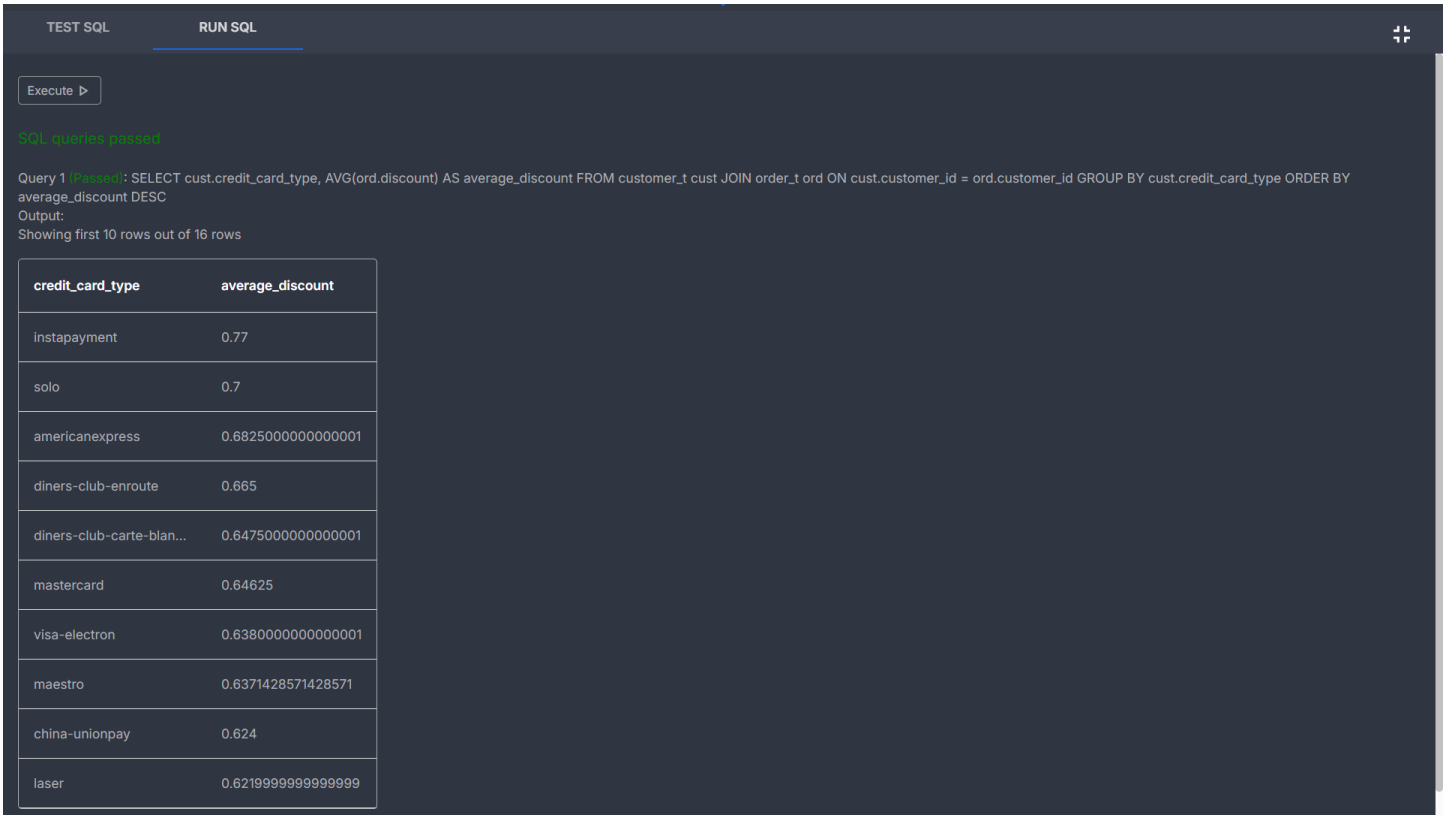
- Sales Performance Concerns: The consistent decline in both revenue and order counts may indicate potential issues in sales performance, such as customer dissatisfaction, increased competition, or market changes.
- Potential for Analysis and Action: These trends highlight the need for further analysis to understand the causes behind the decline in orders and revenue. Investigating customer feedback, market conditions, and pricing strategies could provide insights for improvement.
- Opportunity for Promotions: Given the downward trend in both metrics, the company may consider promotional strategies or marketing campaigns to stimulate interest and boost sales in subsequent quarters.

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

Solution Query:

```
SELECT
   cust.credit_card_type,
   AVG(ord.discount) AS average_discount
FROM
   customer_t cust
JOIN
   order_t ord ON cust.customer_id = ord.customer_id
GROUP BY
   cust.credit_card_type
ORDER BY
   average_discount DESC;
```

Output:



credit_card_type	average_discount
instapayment	0.77
solo	0.7
americanexpress	0.6825000000000001
diners-club-enroute	0.665
diners-club-carte-blan...	0.6475000000000001
mastercard	0.64625
visa-electron	0.6380000000000001
maestro	0.6371428571428571
china-unionpay	0.624
laser	0.6219999999999999

Observations and Insights:

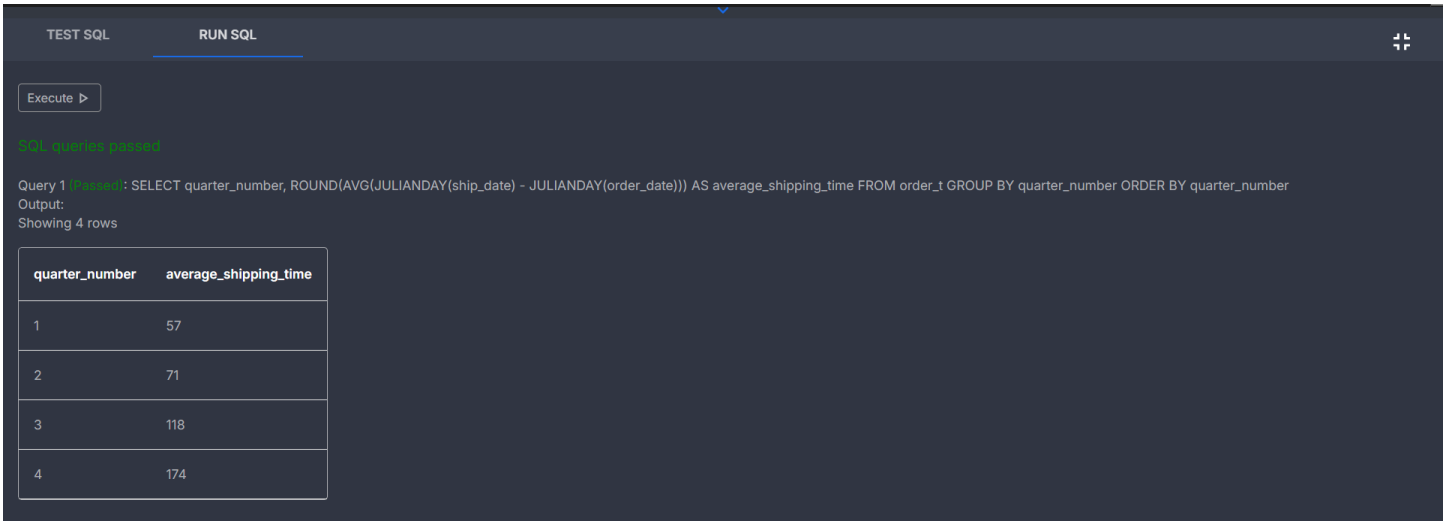
- Instapayment: This credit card type has the highest average discount at 0.77, indicating it might be favored for promotional offers or customer retention strategies.
- Solo: With an average discount of 0.7, Solo card users also benefit from relatively high discounts, suggesting it could be another strategic option for marketing efforts.
- American Express: The average discount for American Express users is 0.6825, showing that this premium card is still providing competitive discounts, appealing to its customer base.
- Diners Club (Enroute and Carte Blanche): Both Diners Club types offer similar average discounts (0.665 and 0.6475, respectively), indicating a consistent approach to rewarding customers using these cards.
- Mastercard: Averaging 0.64625, Mastercard users receive a solid discount, which can enhance loyalty among users of this widely accepted credit card.
- Visa Electron: With an average discount of 0.638, this card type remains competitive, making it a viable option for discount-driven customers.
- Maestro: At 0.6371, the average discount for Maestro cardholders shows a slightly lower discount compared to Visa Electron, yet it still ranks well.
- China UnionPay and Laser: These card types have the lowest average discounts (0.624 and 0.622), suggesting they may not be as prioritized in promotional strategies.
- There is a clear trend that some credit card types, particularly Instapayment and Solo, are more beneficial for customers in terms of discounts, which could be leveraged in targeted marketing campaigns.

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))) AS average_shipping_time
FROM
   order_t
GROUP BY
   quarter_number
ORDER BY
   quarter_number;
```

Output:



Observations and Insights:

- Increasing Trend: There is a noticeable increase in the average shipping time from Quarter 1 to Quarter 4, suggesting that orders are taking significantly longer to ship as the year progresses.
- Quarter 1 Performance: The average shipping time in Quarter 1 is relatively low at 57 days, indicating a more efficient shipping process or fewer delays during this period.
- Quarter 2 to Quarter 3 Spike: The average shipping time jumps from 71 days in Quarter 2 to 118 days in Quarter 3, which may suggest seasonal impacts or operational challenges that arose during the mid-year.
- Significant Delay in Quarter 4: The most substantial increase is seen in Quarter 4, with an average shipping time of 174 days. This could be attributed to holiday seasons, increased demand, or logistical issues during year-end.

- Potential Operational Review Needed: The overall increase in average shipping times throughout the quarters may warrant an operational review to identify bottlenecks and areas for improvement in the shipping process.
- Customer Experience Impact: Longer shipping times can negatively affect customer satisfaction and retention, indicating a need for strategies to improve delivery timelines, especially in the latter half of the year.

Business Metrics Overview

Total Revenue	Total Orders	Total Customers	Average Rating
83115001.10	1000	133	3.135
Last Quarter Revenue	**Last quarter Orders**	**Average Days to Ship**	**% Good Feedback**
15280009.97	199	98	44.1

Business Recommendations

- Target High-Discount Credit Cards: Focus marketing efforts on credit cards that receive higher average discounts, such as Instapayment and Solo, to attract more transactions and boost sales.
- Improve Shipping Efficiency: The average shipping duration has increased to 98 days by the last quarter. Streamlining logistics, reducing delays, and offering faster shipping options will help improve customer experience and satisfaction.
- Address Rising Customer Dissatisfaction: Customer dissatisfaction has increased significantly, with negative feedback rising from 22.26% in Quarter 1 to 59.8% in Quarter 4. Immediate efforts are needed to improve product quality, customer service, and post-sales support.
- Implement Proactive Feedback Mechanisms: Establish a system to collect detailed feedback and act on customer concerns in real-time. This can help identify root causes of dissatisfaction and drive improvements.
- Reverse the Decline in Revenue and Orders: Net revenue and order counts have been steadily declining quarter-over-quarter, with a sharp drop of 22.5% in the last quarter. Strategies such as promotional offers or loyalty programs should be introduced to reinvigorate sales.
- Focus on Customer Retention: With a small pool of repeat customers, efforts should be made to convert one-time buyers into loyal customers through targeted loyalty programs and personalized offers.
- Revise Pricing and Discount Strategies: Regularly review pricing and discounts to maintain competitiveness while managing profit margins. This will be crucial for sustaining growth amid increasing dissatisfaction.
- Enhance Data-Driven Decision Making: Utilize detailed data analytics to forecast future trends and adjust strategies accordingly. This includes monitoring feedback trends and order patterns to anticipate future customer needs.
- Revitalize Quarter 4 Performance: Analyze what caused the sharp drop in Quarter 4 performance and take corrective actions. Replicate strategies from earlier quarters that led to higher revenues and orders.
- Strengthen Customer Relationship Management: Enhance CRM systems to track and understand customer behavior, allowing for more personalized marketing and retention strategies. Better customer management will improve satisfaction and loyalty over time.

Overall, addressing rising customer dissatisfaction, improving operational efficiencies like shipping, and focusing on customer retention will be key to reversing declining trends and driving future growth.