

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:

```
-- Total number of unique customers

SELECT COUNT(DISTINCT CUSTOMER_ID) AS total_customers

FROM order_t;

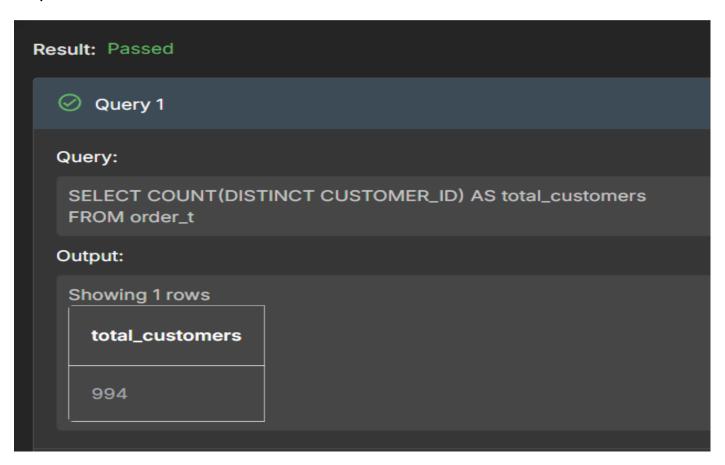
-- Distribution of customers across states

SELECT STATE, COUNT(DISTINCT CUSTOMER_ID) AS customer_count

FROM customer_t

GROUP BY STATE

ORDER BY customer_count DESC;
```





Result: Passed

Query 1

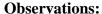
Query:

SELECT STATE, COUNT(DISTINCT CUSTOMER_ID) AS customer_count FROM customer_t GROUP BY STATE ORDER BY customer_count DESC

Output:

Showing first 10 rows out of 37 rows

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





1. Total Customers:

o A total of **994 customers** have placed orders.

2. State-wise Customer Distribution:

- California has the highest customer count with 17 customers, accounting for 1.71% of the total customer base.
- Texas (10 customers, 1.01%) and Florida (9 customers, 0.91%) are the next most significant contributors.
- New York has 7 customers, while Virginia, Michigan, Illinois, and District of Columbia each have 5 customers.
- o Pennsylvania and Ohio follow with 4 customers each.

3. **Top 10 States**:

These states together account for **76 customers**, making up **7.65%** of the total customer base.

- The majority of the customers are distributed across a few states, with California, Texas, and Florida being key contributors. These states present strong opportunities for customer retention and upselling initiatives.
- California, being the state with the highest customer count, should be prioritized for targeted marketing and promotional campaigns.
- Texas and Florida can also be leveraged as growth hubs due to their relatively high customer counts.
- Efforts should be made to increase the customer base in states with fewer than 5 customers to reduce dependency on high-performing states and achieve better geographic balance.

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



customers?

Solution Query:

```
-- Top 5 vehicle makers preferred by customers using INNER JOIN

SELECT
    P.VEHICLE_MAKER,
    COUNT(DISTINCT O.CUSTOMER_ID) AS customer_count

FROM
    ORDER_T O

INNER JOIN
    PRODUCT_T P

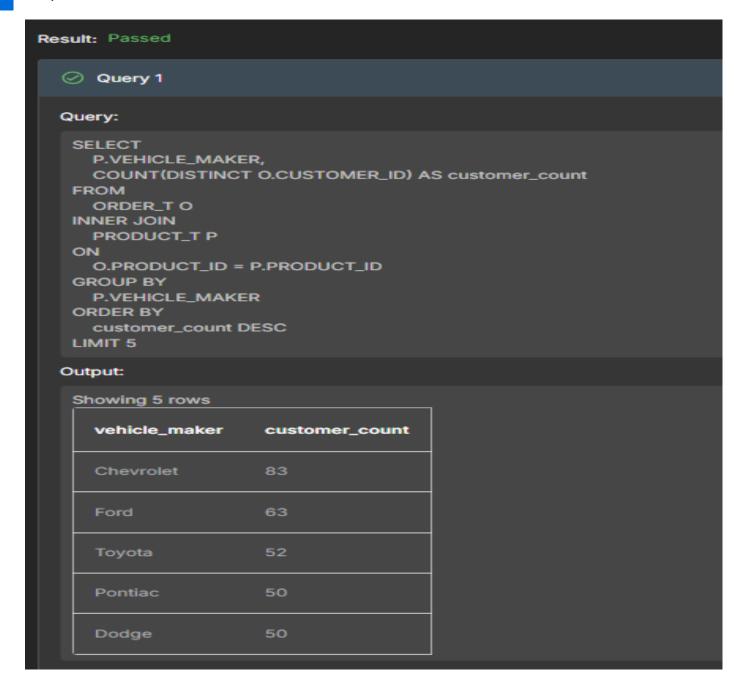
ON
    O.PRODUCT_ID = P.PRODUCT_ID

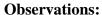
GROUP BY
    P.VEHICLE_MAKER

ORDER BY
    customer_count DESC

LIMIT 5;
```









1. Top Vehicle Makers:

- Chevrolet leads as the most preferred vehicle maker, with 83 unique customers choosing this brand.
- o **Ford** ranks second, with **63 customers**, showing a strong customer base.
- o **Toyota** is the third most popular, preferred by **52 customers**.
- o **Pontiac** and **Dodge** are equally popular, each with **50 customers**.

2. Popularity Distribution:

The top 5 vehicle makers collectively account for a significant portion of customer preferences,
 with a total of 298 customers.

- Chevrolet stands out as the market leader, reflecting its strong brand presence and customer loyalty. Leveraging this preference in marketing campaigns could boost further engagement.
- Ford and Toyota maintain significant popularity, indicating a competitive market presence.
 These brands could benefit from promotions targeting loyal customers to strengthen their position further.
- O Pontiac and Dodge have similar customer counts, suggesting they cater to a niche but stable customer base. Differentiated marketing strategies can help these brands expand their reach.
- The dominance of Chevrolet, Ford, and Toyota indicates that focusing efforts on partnerships or promotions with these brands could maximize customer acquisition and satisfaction.
- The relatively smaller gap between Toyota, Pontiac, and Dodge suggests an opportunity for these brands to compete for market share through targeted campaigns.

Great Learning

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

state?

Solution Query:

```
SELECT *
FROM
(
  SELECT
    state,
    vehicle maker,
    COUNT(c.customer id) AS total customers,
    RANK() OVER (PARTITION BY state ORDER BY COUNT(c.customer_id) DESC) AS
ranking
    FROM product_t p
    JOIN order t o
    on p.product id = o.product id
    JOIN customer_t c
    ON o.customer id = c.customer id
  GROUP BY 1, 2
) AS preferred vehicle
WHERE ranking = 1
ORDER BY 3 DESC;
```



```
Result: Passed
  Query 1
  Query:
   SELECT *
   FROM
        SELECT
            state,
            vehicle_maker,
            COUNT(c.customer_id) AS total_customers,
     RANK() OVER (PARTITION BY state ORDER BY COUNT(c.customer_id) DESC) AS ranking
     FROM product_t p
     JOIN order_t o
     on p.product_id = o.product_id
     JOIN customer_t c
     ON o.customer_id = c.customer_id
        GROUP BY 1, 2
   ) AS preferred_vehicle
   WHERE ranking = 1
   ORDER BY 3 DESC
  Output:
   Showing first 10 rows out of 101 rows
                   vehicle_maker
                                     total_customers
      state
                                                         ranking
      Florida
```



Observations:

- California: Vehicle makers tied for the top spot include Pontiac, Nissan, Ford, and Chevrolet,
 each with two customers.
- o Florida: Volvo and Ford are tied as the most preferred vehicle makers.
- o Indiana: Mazda is the preferred maker.
- Texas: Nissan leads.
- o Alabama: Lincoln and Lexus are tied for the top spot.

- o Nissan appears in both California and Texas, hinting at potential regional strength in these states.
- Luxury brands like Lincoln and Lexus tie for first in Alabama, which may indicate a preference for premium vehicles in this state.



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
 AVG(rating) AS OverallAverageRating
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 rating
    FROM
      order t
  ) AS subquery;
SELECT
  quarter number,
  AVG(rating) AS AverageRatingPerQuarter
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 subquery
GROUP BY
  quarter number
ORDER BY
  quarter number;
```



```
Result: Passed
  Query 1
  Query:
   SELECT
    AVG(rating) AS OverallAverageRating
   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 rating
     FROM
      order_t
    ) AS subquery
  Output:
   Showing 1 rows
     OverallAverageRating
     3.135
```



Result: Passed Query 1 Query: SELECT quarter_number, AVG(rating) AS AverageRatingPerQuarter 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 subquery **GROUP BY** quarter_number ORDER BY quarter_number Output: Showing 4 rows AverageRatingPerQua... quarter_number 3.554838709677419 3.354961832061069 2.9563318777292578 4 2.3969849246231156



Observations:

1. Overall Average Rating:

The overall average customer rating across all orders is **3.135**, indicating that customer feedback tends to fall between "Okay" (rating 3) and "Good" (rating 4).

2. Quarter-wise Average Ratings:

- O Quarter 1: The average rating is 3.55, which is the highest among all quarters. This suggests relatively better customer satisfaction during this period, closer to "Good."
- Quarter 2: The average rating drops slightly to 3.35, indicating a slight decline in customer satisfaction.
- Quarter 3: The rating decreases further to 2.96, approaching "Okay."
- Quarter 4: The lowest average rating is 2.40, which indicates dissatisfaction, closer to "Bad."

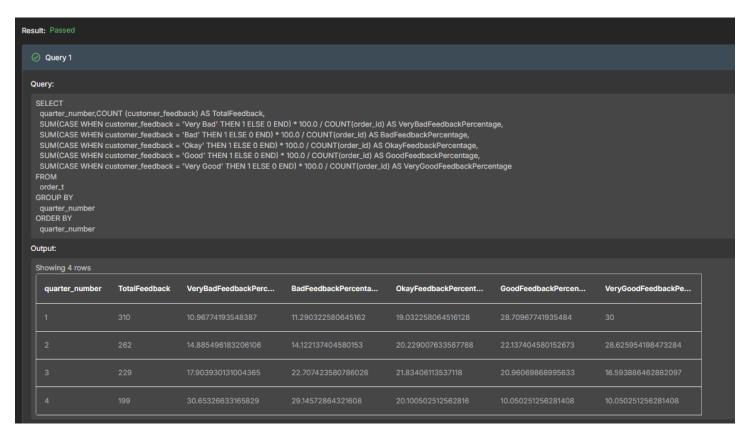
- The data suggests a clear downward trend in customer ratings over the quarters. Quarter 1 has the highest satisfaction, while Quarter 4 has the lowest. This may indicate seasonal issues affecting customer experience, such as high demand, product availability, or service delays towards the end of the year.
- The significant drop in Quarter 4 could be due to challenges like holiday rushes, increased order volumes, or reduced service quality. Addressing these factors—such as improving logistics or customer support during peak seasons—could help boost ratings in this period.
- An overall average rating of 3.135 indicates a moderate level of satisfaction. Efforts to enhance service or product quality could push more feedback into the "Good" and "Very Good" categories, improving the overall score.

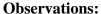


Question 5: Find the percentage distribution of feedback from the customers. Are customers getting more dissatisfied over time?

Solution Query:

```
SELECT
 quarter number, COUNT (customer feedback) AS TotalFeedback,
  SUM(CASE WHEN customer feedback = 'Very Bad' THEN 1 ELSE 0 END) * 100.0 /
COUNT (order id) AS VeryBadFeedbackPercentage,
  SUM(CASE WHEN customer feedback = 'Bad' THEN 1 ELSE 0 END) * 100.0 /
COUNT (order id) AS BadFeedbackPercentage,
  SUM(CASE WHEN customer feedback = 'Okay' THEN 1 ELSE 0 END) * 100.0 /
COUNT (order id) AS OkayFeedbackPercentage,
  SUM(CASE WHEN customer feedback = 'Good' THEN 1 ELSE 0 END) * 100.0 /
COUNT (order id) AS GoodFeedbackPercentage,
  SUM(CASE WHEN customer_feedback = 'Very Good' THEN 1 ELSE 0 END) * 100.0 /
COUNT(order id) AS VeryGoodFeedbackPercentage
FROM
 order t
GROUP BY
 quarter number
ORDER BY
 quarter number;
```







1. Feedback Distribution Across Quarters:

Ouarter 1:

- Positive feedback (Good + Very Good) accounts for **58.71%**.
- Negative feedback (Very Bad + Bad) is 22.26%.

Quarter 2:

• Positive feedback declines to **50.76%**, while negative feedback increases to **29.01%**.

Quarter 3:

Positive feedback drops significantly to 37.55%, and negative feedback rises to 40.61%.

Quarter 4:

Positive feedback is at its lowest at 20.10%, while negative feedback reaches 59.80%.

2. Trend in Dissatisfaction:

- A clear upward trend in dissatisfaction is observed over time. In Quarter 1, only 22.26% of customers reported negative feedback, but by Quarter 4, this has increased to almost 60%.
- The percentage of Very Bad feedback increases steadily, from 10.97% in Quarter 1 to 30.65% in Ouarter 4.
- Conversely, the percentage of Very Good feedback drops from 30% in Quarter 1 to 10.05% in Quarter 4.

Insights:

1. Rising Customer Dissatisfaction:

The data clearly shows a **deterioration in customer satisfaction** over the year. By Quarter 4, a majority of customers are reporting negative experiences, with over 30% giving "Very Bad" feedback.

2. Potential Causes:

The rising dissatisfaction could be linked to operational issues, seasonal peaks, resource constraints, or changes in product/service quality. For example, Quarter 4 might face challenges such as holiday season stress, shipping delays, or product shortages.

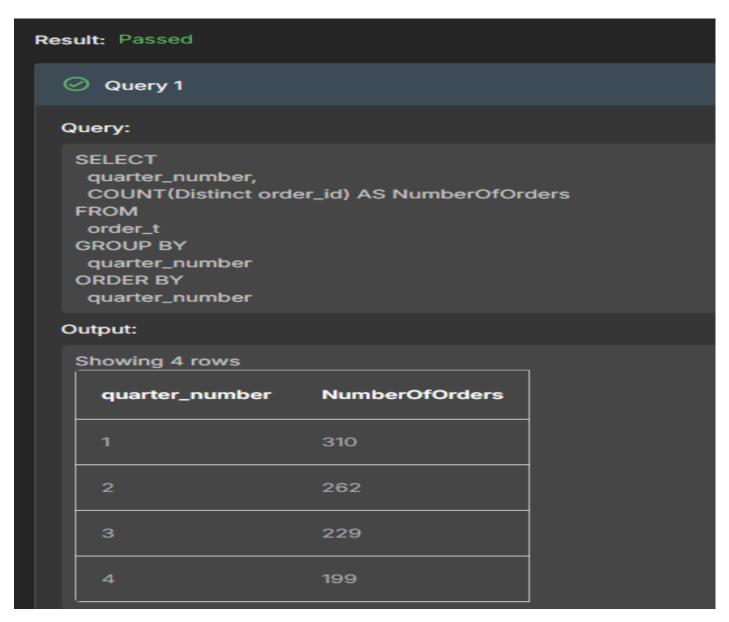






Solution Query:

```
SELECT
  quarter_number,
  COUNT(Distinct order_id) AS NumberOfOrders
FROM
  order_t
GROUP BY
  quarter_number
ORDER BY
  quarter number;
```





Observations:

1. Quarterly Order Trend:

- o Quarter 1: The highest number of orders, with 310.
- Quarter 2: A decline to 262 orders.
- o Quarter 3: Further decline to 229 orders.
- Quarter 4: The lowest number of orders, with only 199.

2. Overall Trend:

 The data shows a steady decline in the number of orders across all quarters. There is a clear downward trajectory from Quarter 1 through Quarter 4.

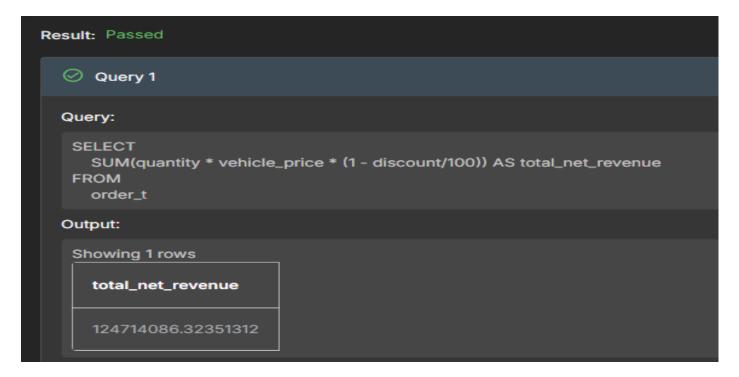
- The consistent decrease in order volume suggests potential seasonal factors, reduced customer demand, or operational constraints impacting order processing.
- The decline in orders may be correlated with the increasing dissatisfaction observed in customer feedback. As customer satisfaction drops, it may contribute to fewer repeat orders or reduced customer loyalty.



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

Solution Query:

```
SELECT
    SUM(quantity * vehicle price * (1 - discount/100)) AS total net revenue
FROM
    order t;
SELECT
    quarter number,
    SUM(quantity * vehicle price * (1 - discount/100)) AS net revenue,
    ROUND (
        (SUM(quantity * vehicle_price * (1 - discount/100)) -
        LAG(SUM(quantity * vehicle price * (1 - discount/100)))
        OVER (ORDER BY quarter number)) * 100.0 /
        LAG(SUM(quantity * vehicle_price * (1 - discount/100)))
        OVER (ORDER BY quarter number),
        2
    ) AS gog percentage change
FROM
   order_t
GROUP BY
    quarter_number
ORDER BY
    quarter number;
```





Result: Passed



Query 1

Query:

```
SELECT
  quarter_number,
  SUM(quantity * vehicle_price * (1 - discount/100)) AS net_revenue,
  ROUND(
    (SUM(quantity * vehicle_price * (1 - discount/100)) -
    LAG(SUM(quantity * vehicle_price * (1 - discount/100)))
    OVER (ORDER BY quarter_number)) * 100.0 /
    LAG(SUM(quantity * vehicle_price * (1 - discount/100)))
    OVER (ORDER BY quarter_number),
    2
  ) AS qoq_percentage_change
FROM
  order_t
GROUP BY
  quarter_number
ORDER BY
  quarter_number
```

Output:

Showing 4 rows

quarter_number	net_revenue	qoq_percentage_change
1	39421580.15929598	
2	32715830.33996199	-17.01
3	29229896.19364898	-10.66
4	23346779.63060599	-20.13



Observations:

1. Total Net Revenue:

o The company generated a total net revenue of \$124,714,086.32 over the analyzed period.

2. Quarterly Net Revenue:

- O Quarter 1: Net revenue was \$39,421,580.16, the highest of all quarters.
- O Quarter 2: Revenue dropped to \$32,715,830.34, showing a -17.01% decline from Quarter 1.
- Quarter 3: Revenue further declined to \$29,229,896.19, with a -10.66% decrease compared to Quarter 2.
- Quarter 4: The lowest revenue was recorded at \$23,346,779.63, representing a -20.13% decline from Quarter 3.

- The data reveals a **consistent quarter-over-quarter decline** in net revenue. The largest decline occurred in Quarter 4, with a significant **-20.13%** drop compared to Quarter 3. This mirrors the trend in customer dissatisfaction and declining order volumes.
- o The downward trend in net revenue correlates with:
 - A decline in order volume across quarters.
 - Increasing customer dissatisfaction, as indicated by rising negative feedback percentages.
- Possible factors include reduced demand, increased cancellations, higher discounting, or operational inefficiencies affecting sales.

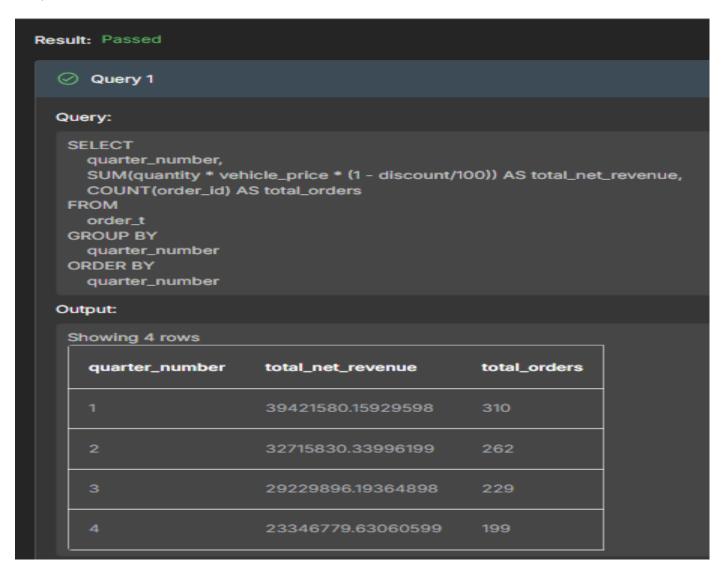


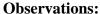
Question 8: What is the trend of net revenue and orders by

quarters?

Solution Query:

```
SELECT
    quarter_number,
    SUM(quantity * vehicle_price * (1 - discount/100)) AS total_net_revenue,
    COUNT(order_id) AS total_orders
FROM
    order_t
GROUP BY
    quarter_number
ORDER BY
    quarter_number;
```







1. Consistent Decline in Net Revenue and Orders:

Net Revenue:

- Declined steadily across all quarters, with the largest drop in Quarter 4 (-20.13%).
- The cumulative decrease from Quarter 1 to Quarter 4 is significant, indicating a sustained decline in financial performance.

Total Orders:

- Similarly, order volume decreased each quarter, with a total drop of approximately **35.8%** from Quarter 1 (310 orders) to Quarter 4 (199 orders).
- The decline in revenue closely follows the trend in the number of orders, indicating that reduced order volume is a primary driver of lower revenues.
- o This suggests **customer demand** or operational challenges might be affecting sales.

- The simultaneous decrease in orders and revenue suggests a **drop in customer demand** over time, possibly due to seasonal changes, market saturation, or customer dissatisfaction.
- The presence of discounts (as inferred from revenue calculation) could be contributing to the revenue decline if larger discounts are offered in later quarters without driving higher order volumes.
- Factors like economic conditions, supply chain disruptions, or internal inefficiencies may also be influencing this trend.



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

Solution Query:

```
SELECT
    T1.credit_card_type,
    AVG(T2.discount) AS AverageDiscount
FROM
    customer_t T1
INNER JOIN
    order_t T2 ON T1.customer_id = T2.customer_id
GROUP BY
    T1.credit_card_type
ORDER BY
    T1.credit_card_type;
```



Result: Passed Query 1 Query: SELECT T1.credit_card_type, AVG(T2.discount) AS AverageDiscount customer_t T1 INNER JOIN order_t T2 ON T1.customer_id = T2.customer_id **GROUP BY** T1.credit_card_type ORDER BY T1.credit_card_type Output: Showing first 10 rows out of 16 rows credit_card_type AverageDiscount 0.68250000000000001 americanexpress bankcard 0.5933333333333334 0.624china-unionpay diners-club-carte-blan... 0.64750000000000001 diners-club-enroute 0.665diners-club-international 0.6diners-club-us-ca 0.5instapayment 0.6083050847457627 jeb 0.621999999999999



Observations:

1. Highest Average Discount:

o **Instapayment** credit cards receive the highest average discount of **0.77%**.

2. Moderate Average Discounts:

Credit card types like American Express (0.68%), Diners Club Enroute (0.67%), and Diners
 Club Carte Blanche (0.65%) also enjoy relatively higher discounts.

3. Lowest Average Discount:

o **Diners Club US/CA** offers the lowest average discount at **0.5%**.

4. Credit Cards with Comparable Discounts:

Many credit card types, such as **China UnionPay** (0.62%), **Laser** (0.62%), and **JCB** (0.61%), have fairly similar average discount rates.

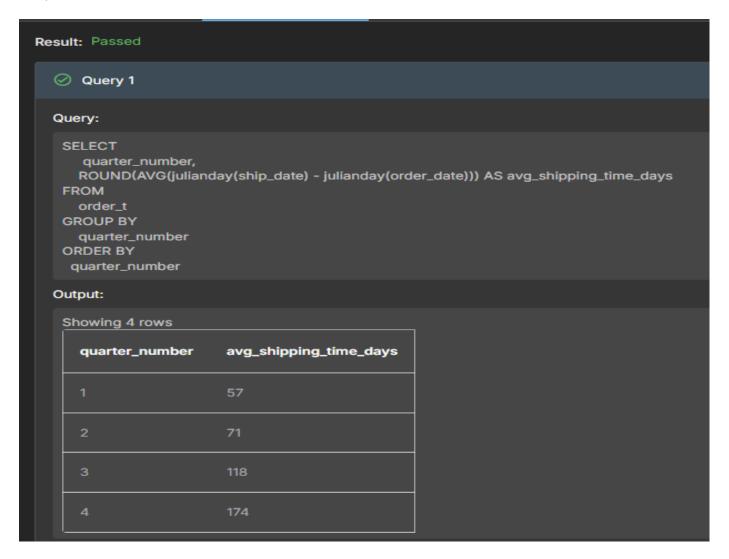
- The average discount does not vary significantly across credit card types, with most falling between 0.6% and 0.77%.
- The slight differences in discount percentages might reflect promotional partnerships or customer segmentation strategies based on credit card usage.
- Higher discounts for certain card types (like Instapayment) could indicate targeted promotions,
 aiming to drive more transactions with these payment methods.



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
avg_shipping_time_days
FROM
    order_t
GROUP BY
    quarter_number
ORDER BY
    quarter_number;
```





Observations:

1. Increasing Shipping Delays:

- The average shipping time increases significantly from Quarter 1 (57 days) to Quarter 4 (174 days).
- The delays escalate each quarter, indicating a consistent decline in operational efficiency or external supply chain challenges.

2. Largest Quarterly Jump:

- The biggest increase in shipping time occurs between Quarter 3 (118 days) and Quarter 4 (174 days), suggesting potential issues such as:
 - Supply chain disruptions.
 - Increased order volumes without adequate capacity.
 - Workforce shortages or logistics bottlenecks.

- The rising shipping time could be contributing to customer dissatisfaction (as indicated by declining feedback scores).
- There may be inefficiencies in order processing, inventory management, or logistics, which require immediate attention.

Business Metrics Overview



Total Revenue	Total Orders	Total Customers	Average Rating
124714086.32351312	1000	994	3.135
Last Quarter Revenue	Last quarter Orders	Average Days to Ship	% Good Feedback
23346779.63060599	199	97.964	21.5

Business Recommendations

1. Improve Customer Satisfaction (Average Rating and Feedback)

• Enhance Product/Service Quality:

The average rating of 3.135 and only 21.5% good feedback indicate low customer satisfaction. Conduct root cause analysis by reviewing product reviews and complaints. Address quality issues or service-related challenges to improve ratings.

• Focus on After-Sales Support:

Implement robust customer support to address issues promptly. Offer satisfaction guarantees or refunds to regain customer trust.

• Encourage Positive Feedback:

Actively request feedback from satisfied customers to boost the percentage of good reviews. Incentivize feedback by offering small discounts or loyalty points for completing surveys.

2. Increase Revenue and Orders

• Promote Customer Retention:

With 994 customers and 1000 total orders, many customers are single-time buyers. Develop loyalty programs to encourage repeat purchases. Send personalized offers based on past purchase behavior.

• Upsell and Cross-Sell:

Encourage customers to spend more by recommending complementary products or services at checkout. Use targeted marketing campaigns based on customer preferences.

• Boost Marketing Campaigns:

Invest in digital ads, email campaigns, and social media promotions to increase visibility and attract new customers. Focus on high-revenue products or best-sellers to maximize returns.

3. Optimize Shipping and Fulfillment



• Reduce Shipping Times:

The average days to ship is 97.96, which is unacceptably high and likely a key contributor to customer dissatisfaction. Streamline supply chain processes, renegotiate with shipping partners, or explore faster delivery options.

• Set Accurate Expectations:

Clearly communicate shipping timelines to customers. If faster shipping isn't feasible, provide proactive updates to manage expectations and avoid negative feedback.

4. Boost Quarterly Revenue

• Focus on Seasonality and Trends:

Last quarter's revenue (\$23.35M) is lower compared to total revenue, suggesting possible seasonality or declining sales. Identify trends and focus on key periods to increase sales through targeted promotions.

• Introduce Limited-Time Offers:

Run time-bound discounts or bundle deals to increase last-quarter orders (currently 199). Create urgency through exclusive campaigns to boost immediate purchases.

5. Competitor Benchmarking and Innovation

• Study Competitors' Strategies:

Analyze competitors' shipping policies, pricing models, and customer satisfaction initiatives. Adopt best practices to stay competitive.

• Innovate Product Offerings:

Introduce new or improved products based on market trends and customer preferences. Diversify offerings to appeal to a broader audience.