
PGP-DSBA PROJECT REPORT

SQL – Coded Project

by Ishaan Shakti Jayaraman

PGPDSBA.OJULY24.A



Problem Statements 4

Business Questions 5

 Question 1: Find the total number of customers who have placed orders. What is the distribution of the customers across states?..... 5

 Question 2: Which are the top 5 vehicle makers preferred by the customers? 7

 Question 3: Which is the most preferred vehicle maker in each state? 8

 Question 4: Find the overall average rating given by the customers. What is the average rating in each quarter?..... 9

 Question 5: Find the percentage distribution of feedback from the customers. Are customers getting more dissatisfied over time?..... 11

 Question 6: What is the trend of the number of orders by quarter? 12

 Question 7: Calculate the net revenue generated by the company. What is the quarter-over-quarter % change in net revenue? 13

 Question 8: What is the trend of net revenue and orders by quarters? 14

 Question 9: What is the average discount offered for different types of credit cards? 15

 Question 10: What is the average time taken to ship the placed orders for each quarter?..... 16

Business Metrics Overview 17

Business Recommendations 17

Table of Figures

Figure 1 - Output of total customers.....	5
Figure 2 - Output of customers by state	6
Figure 3 - Output of top vehicle makers by customers	7
Figure 4 - Output of most preferred vehicle by state	8
Figure 5 - Output of overall average rating.....	10
Figure 6 - Output of average rating per quarter.....	10
Figure 7 - Output of feedback distribution per quarter	11
Figure 8 - Output of number of orders per quarter	12
Figure 9 - Output of revenue by quarter and percentage change.....	13
Figure 10 - Output of net revenue & total orders per quarter.....	14
Figure 11 - Output of average discount by credit card type	15
Figure 12 - Output of average shipping time per quarter.....	16

New Wheels Project

Introduction to SQL

Problem Statements

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.

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 customer_id) AS total_customers  
FROM order_t;
```

```
SELECT state, COUNT(DISTINCT customer_id) AS number_of_customers  
FROM customer_t  
GROUP BY state  
ORDER BY number_of_customers DESC;
```

Output:

Figure 1 - Output of total customers

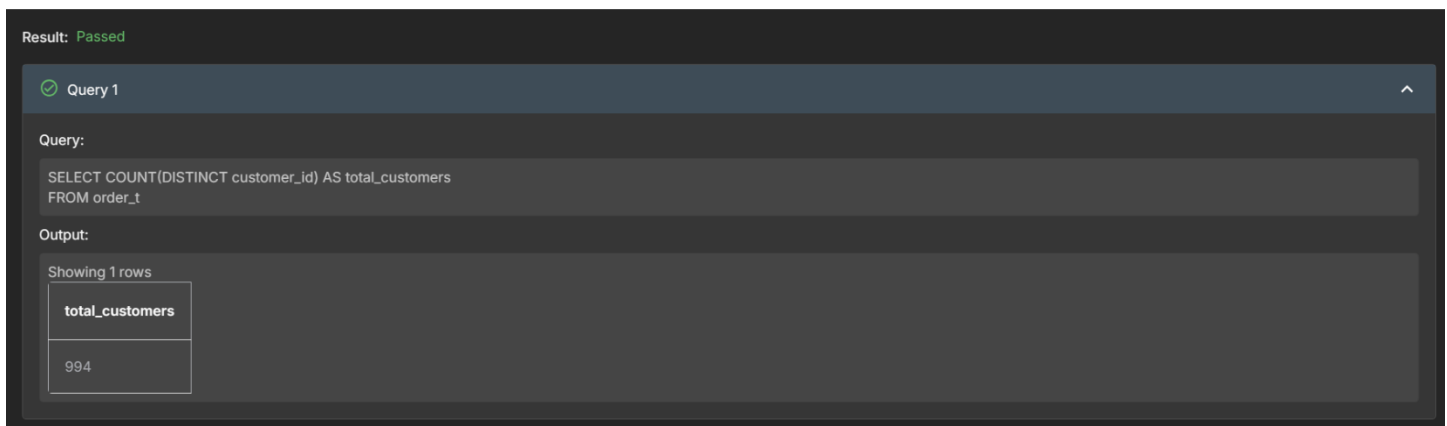
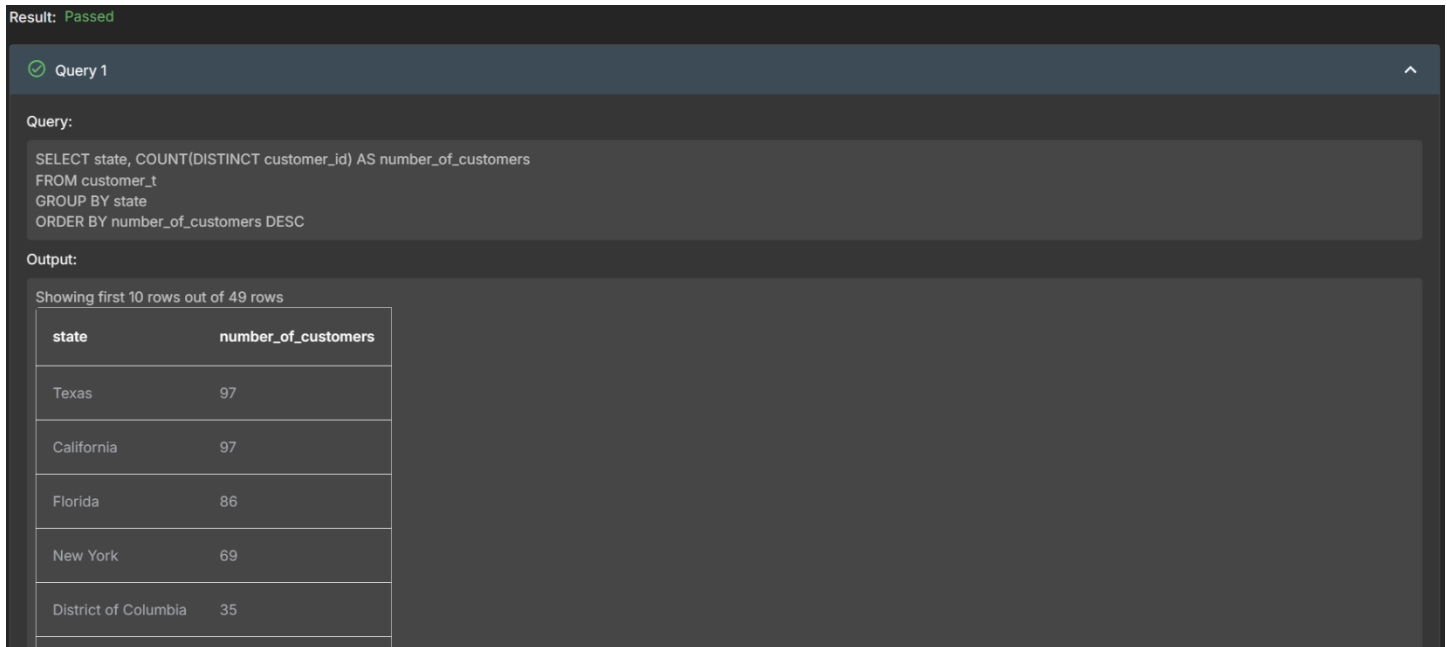


Figure 2 - Output of customers by state



Observations and Insights:

- There are a total of 994 unique customers in the database who have placed orders.
- Texas, California, Florida, New York and District of Columbia are the top 5 states of customers with around 384 customers.
- Maine, Vermont, Wyoming, Mississippi and North Dakota have the lowest number of customers.

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

Solution Query:

```
SELECT
    vehicle_maker AS top_vehicle_makers,
    COUNT(customer_id) AS total_customers
FROM product_t JOIN customer_t
GROUP BY 1
ORDER BY 2 DESC
LIMIT 5;
```

Output:

Figure 3 - Output of top vehicle makers by customers

Result: Passed

Query 1

Query:

```
SELECT
    vehicle_maker AS top_vehicle_makers,
    COUNT(customer_id) AS total_customers
FROM product_t JOIN customer_t
GROUP BY 1
ORDER BY 2 DESC
LIMIT 5
```

Output:

Showing 5 rows

top_vehicle_makers	total_customers
Chevrolet	82502
Ford	62622
Toyota	51688
Pontiac	49700
Dodge	49700

Observations and Insights:

- Chevrolet, Ford, Toyota, Pontiac and Dodge are the top 5 car makers preferred by the customers
- Chevrolet has the largest number of customers, followed by Ford and Toyota.
- Pontiac and Dodge have similar number of customers of around 49,700.

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

Solution Query:

```
SELECT state, vehicle_maker
FROM (
    SELECT
        state,
        vehicle_maker,
        COUNT(c.customer_id) AS customer_count,
        RANK() OVER (PARTITION BY c.state ORDER BY COUNT(c.customer_id)
        DESC) AS maker_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 RankedMakers
WHERE maker_rank = 1
ORDER BY state;
```

Output:

Figure 4 - Output of most preferred vehicle by state

Result: Passed

Query 1

Query:

```
SELECT state, vehicle_maker
FROM (
    SELECT
        state,
        vehicle_maker,
        COUNT(c.customer_id) AS customer_count,
        RANK() OVER (PARTITION BY c.state ORDER BY COUNT(c.customer_id) DESC) AS maker_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 RankedMakers
WHERE maker_rank = 1
ORDER BY state
```

Output:

Showing first 10 rows out of 143 rows

state	vehicle_maker
Alabama	Dodge
Alaska	Chevrolet
Arizona	Pontiac
Arizona	Cadillac
Arkansas	Volkswagen

Observations and Insights:

- Chevrolet has a considerable number of orders so it is most preferred by various states in the country.
- Some states have more than one preferred vehicle maker e.g Arizona where both Pontiac and Cadillac are preferred.
- Dodge, Pontiac and Ford are among the most preferred vehicle in majority of the states.

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(  
    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 overall_average_rating  
FROM order_t;  
  
SELECT quarter_number,  
    ROUND(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  
    ),2) AS average_rating  
FROM order_t  
GROUP BY 1  
ORDER BY 1 ASC;
```

Output:

Figure 5 - Output of overall average rating

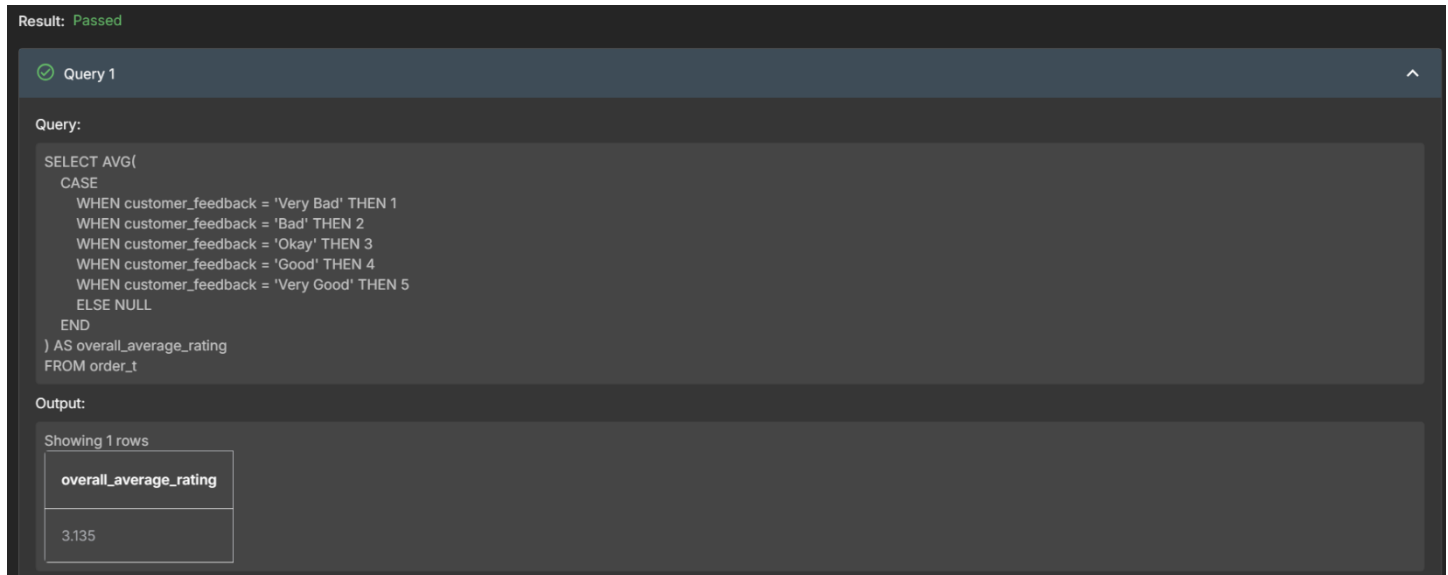
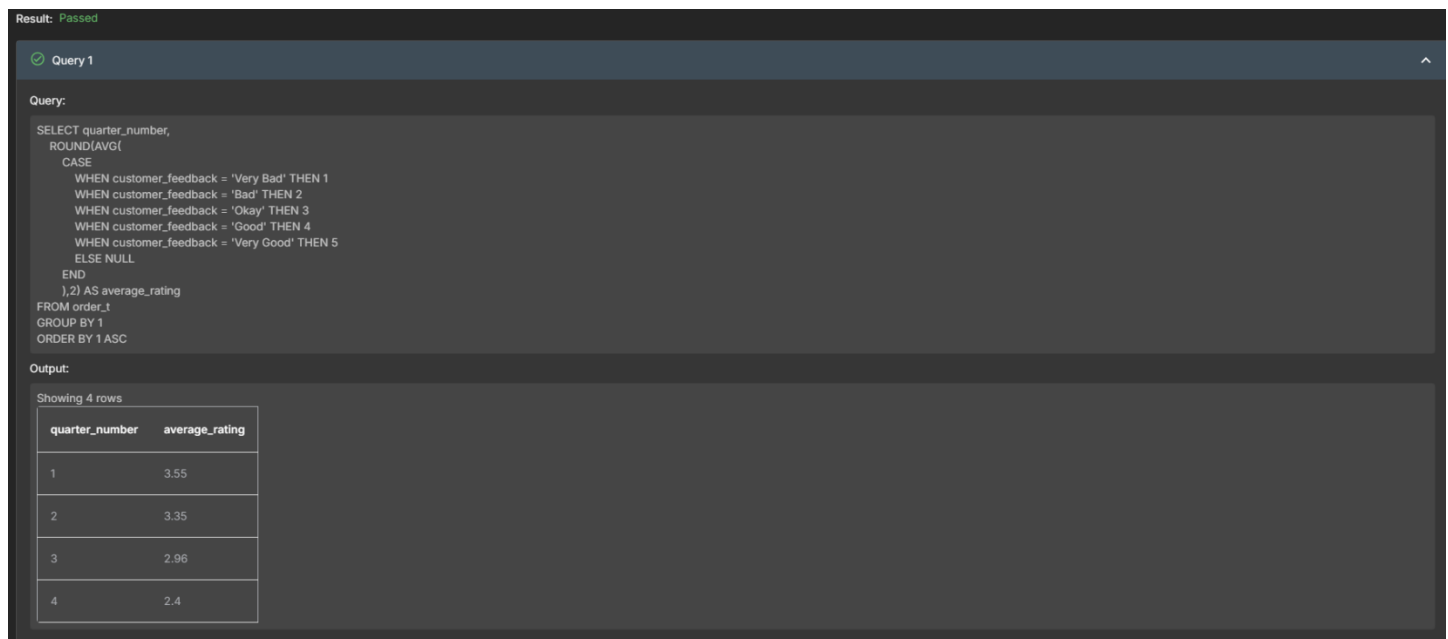


Figure 6 - Output of average rating per quarter



Observations and Insights:

- The overall average rating is 3.135 out of 5.
- The average rating in the first quarter is 3.55 and the average rating in the last quarter is 2.4
- There is an overall decrease in ratings from the customers over the quarters.
- This shows a decline in customer satisfaction throughout the year.

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

Solution Query:

```
SELECT quarter_number,
       ROUND(COUNT(CASE WHEN customer_feedback = 'Very Bad' THEN 1 END) *
100.0 / COUNT(customer_feedback), 2) AS very_bad,
       ROUND(COUNT(CASE WHEN customer_feedback = 'Bad' THEN 1 END) * 100.0 /
COUNT(customer_feedback), 2) AS bad,
       ROUND(COUNT(CASE WHEN customer_feedback = 'Okay' THEN 1 END) * 100.0
/ COUNT(customer_feedback), 2) AS okay,
       ROUND(COUNT(CASE WHEN customer_feedback = 'Good' THEN 1 END) * 100.0
/ COUNT(customer_feedback), 2) AS good,
       ROUND(COUNT(CASE WHEN customer_feedback = 'Very Good' THEN 1 END) *
100.0 / COUNT(customer_feedback), 2) AS very_good
FROM order_t
GROUP BY quarter_number
ORDER BY quarter_number;
```

Output:

Figure 7 - Output of feedback distribution per quarter

Result: Passed

Query 1

Query:

```
SELECT quarter_number,
       ROUND(COUNT(CASE WHEN customer_feedback = 'Very Bad' THEN 1 END) * 100.0 / COUNT(customer_feedback), 2) AS very_bad,
       ROUND(COUNT(CASE WHEN customer_feedback = 'Bad' THEN 1 END) * 100.0 / COUNT(customer_feedback), 2) AS bad,
       ROUND(COUNT(CASE WHEN customer_feedback = 'Okay' THEN 1 END) * 100.0 / COUNT(customer_feedback), 2) AS okay,
       ROUND(COUNT(CASE WHEN customer_feedback = 'Good' THEN 1 END) * 100.0 / COUNT(customer_feedback), 2) AS good,
       ROUND(COUNT(CASE WHEN customer_feedback = 'Very Good' THEN 1 END) * 100.0 / COUNT(customer_feedback), 2) AS very_good_percentage
FROM order_t
GROUP BY quarter_number
ORDER BY quarter_number
```

Output:

Showing 4 rows

quarter_number	very_bad	bad	okay	good	very_good_percentage
1	10.97	11.29	19.03	28.71	30
2	14.89	14.12	20.23	22.14	28.63
3	17.9	22.71	21.83	20.96	16.59
4	30.65	29.15	20.1	10.05	10.05

Observations and Insights:

- During the first 2 quarters the percentage of good and very good ratings form around 60% of total customer ratings
- During the last 2 quarters, there is a decline showing that majority of customer rating move lower towards very bad and bad.
- Overall, There a decline in ratings throughout the quarters.

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

Solution Query:

```
SELECT quarter_number, COUNT(order_id) AS number_of_orders
FROM order_t
GROUP BY quarter_number
ORDER BY quarter_number;
```

Output:

Figure 8 - Output of number of orders per quarter

Result: Passed

Query 1

Query:

```
SELECT quarter_number, COUNT(order_id) AS number_of_orders
FROM order_t
GROUP BY quarter_number
ORDER BY quarter_number
```

Output:

Showing 4 rows

quarter_number	number_of_orders
1	310
2	262
3	229
4	199

Observations and Insights:

- The highest number of quarters are in quarter 1 with 310 orders.
- The 2nd quarter has a total of 262 orders.
- 229 orders were made by customers in the 3rd quarter.
- The last quarter had a total of 199 orders.
- Overall, there is a decline in orders by customers throughout the quarters.

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(
         SUM(quantity * vehicle_price * (1 - discount)), 2
       ) AS quarterly_revenue,
       ROUND(
         (
           SUM(quantity * vehicle_price * (1 - discount))
           - LAG(SUM(quantity * vehicle_price * (1 - discount)), 1, 0)
           OVER (ORDER BY quarter_number)
         ) * 100 /
         LAG(SUM(quantity * vehicle_price * (1 - discount)), 1, 0)
         OVER (ORDER BY quarter_number), 2
       ) AS qoq_percentage_change
FROM order_t
GROUP BY quarter_number
ORDER BY quarter_number;
```

Output:

Figure 9 - Output of revenue by quarter and percentage change

Result: Passed

Query 1

Query:

```
SELECT quarter_number,
       ROUND(
         SUM(quantity * vehicle_price * (1 - discount)), 2
       ) AS quarterly_revenue,
       ROUND(
         (
           SUM(quantity * vehicle_price * (1 - discount))
           - LAG(SUM(quantity * vehicle_price * (1 - discount)), 1, 0)
           OVER (ORDER BY quarter_number)
         ) * 100 /
         LAG(SUM(quantity * vehicle_price * (1 - discount)), 1, 0)
         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	quarterly_revenue	qoq_percentage_change
1	18032549.9	
2	13122995.76	-27.23
3	8882298.84	-32.32
4	8573149.28	-3.48

Observations and Insights:

- The first quarter shows the highest revenue earned at around 18,032,550.
- There is a decline by about 27% in the second quarter.
- A further decline of 32.32 percent in revenue from the 2nd to the 3rd quarter.
- The last quarter has the lowest revenue earned with a decline by 3.48% from the previous quarter.
- Overall, there is a significant decrease in revenue over the quarters.

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 revenue,
       COUNT(order_id) AS total_order
FROM order_t
GROUP BY quarter_number
ORDER BY quarter_number;
```

Output:

Figure 10 - Output of net revenue & total orders per quarter

Result: Passed

Query 1

Query:

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

Output:

Showing 4 rows

quarter_number	revenue	total_order
1	18032549.9	310
2	13122995.76	262
3	8882298.84	229
4	8573149.28	199

Observations and Insights:

- The revenue has declined from 18032549.9 in quarter 1 to 8573149.28 in quarter 4
- The orders have declined from 310 orders in quarter 1 to 199 orders in quarter 4
- Overall, there is a decline in both revenue and total orders.
- Revenue and total orders are correlated as with more orders, the total revenue also increases.

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

Solution Query:

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

Output:

Figure 11 - Output of average discount by credit card type

Result: Passed

Query 1

Query:

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

Output:

Showing first 10 rows out of 16 rows

credit_card_type	average_discount
laser	0.64
mastercard	0.63
visa-electron	0.62
maestro	0.62
instapayment	0.62
china-unionpay	0.62

Observations and Insights:

- The highest discount rate is offered for laser cards at 0.64.
- Mastercard credit card is offered a discount rate of 0.63.
- Visa-electron, maestro and instapayment and china-unionpay is offered a rate of 0.62.
- There are a total of 16 credit card types that are offered a discount.

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

Output:

Figure 12 - Output of average shipping time per quarter

Result: Passed

Query 1

Query:

```
SELECT quarter_number,
       ROUND(AVG(julianday(ship_date) - julianday(order_date)), 2) AS average_shipping_time
FROM order_t o
GROUP BY quarter_number
ORDER BY quarter_number
```

Output:

Showing 4 rows

quarter_number	average_shipping_time
1	57.17
2	71.11
3	117.76
4	174.1

Observations and Insights:

- The first quarter has an average shipping time of 57 days.
- Average shipping time is 71 days for the 2nd quarter.
- The 3rd quarter has an average shipping time of 118 days.
- The last quarter has an average shipping time of 174 days.
- Overall, there is an increase of shipping time to deliver the product to the customer, this could explain the overall decrease in customer satisfaction.

Business Metrics Overview

Total Revenue	Total Orders	Total Customers	Average Rating
48,610,993.78	1000	994	3.14
Last Quarter Revenue	Last quarter Orders	Average Days to Ship	% Good Feedback
8,573,149.28	199	97.96	21.5%

Business Recommendations

- The total orders have declined from 310 in Q1 to 199 in Q4, with a total of 1000 orders, market research should be conducted to identify the main cause of the decline in orders.
- Customer ratings dropped significantly from Q1 to Q4, indicating growing customer dissatisfaction, Surveys can be conducted to identify the key reasons why there is a shift to negative reviews.
- Chevrolet and Ford are the most preferred vehicle maker among customers, marketing and promotional efforts can be designed around those brands to maintain customer interest.
- The average delivery time across quarters has been increasing significantly which may impact customer satisfaction. The process must be analyzed to identify why there is an increase and it should be streamlined to reduce the delivery time.
- Declining orders have led to reduced revenue and a lack of profitability, strategies should be implemented such as personalized offers, promotions, and loyalty programs to increase orders.
- Conducting market research to identify if there is a reason outside of company control which impacts orders such as rising petrol costs, pandemic emergencies or new competitors etc.
- Additional marketing and promotion efforts can be done specifically to the top states by orders such as Texas, California, Florida, New York and District of Columbia and more.
- Marketing efforts can be made for each state based on the most favored brand of cars.
- Offering high discounts could be impacting the revenue considering that satisfaction has overall declined these discounts have not helped improving it, an analysis should be conducted to identify the overall impact of discounts on customer satisfaction.
- There are over 994 customers across multiple states, marketing efforts can be tailored to improve number of orders in regions with low sales orders.