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.

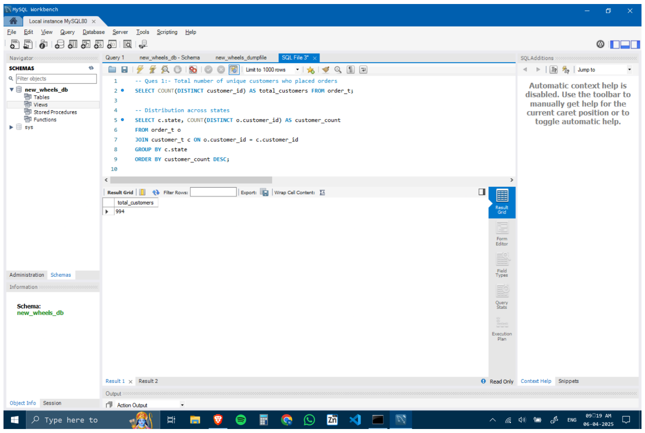
Business Questions

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

Solution Query:

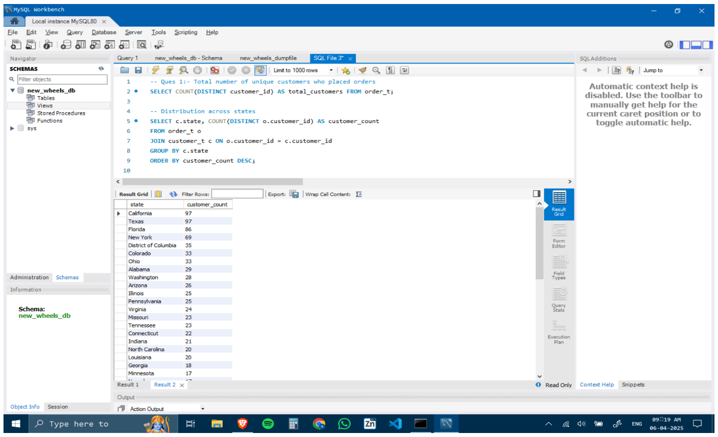
```
-- Total number of unique customers who placed orders

SELECT COUNT(DISTINCT customer_id) AS total_customers FROM order_t;

-- Distribution across states

SELECT c.state, COUNT(DISTINCT o.customer_id) AS customer_count

FROM order_t o

JOIN customer_t c ON o.customer_id = c.customer_id

GROUP BY c.state

ORDER BY customer_count DESC;
```

Output:



```
-- Ques 1:- Total number of unique customers who placed orders
SELECT COUNT(DISTINCT customer_id) AS total_customers FROM order_t;

-- Distribution across states
SELECT c.state, COUNT(DISTINCT o.customer_id) AS customer_count
FROM order_t o
JOIN customer_t c ON o.customer_id = c.customer_id
GROUP BY c.state
ORDER BY customer_count DESC;
```

total_customers
994



state	customer_count
California	97
Texas	97
Florida	86
New York	69
District of Columbia	35
Colorado	33
Ohio	33
Alabama	29
Washington	28
Arizona	26
Illinois	25
Pennsylvania	25
Virginia	24
Missouri	23
Tennessee	23
Connecticut	22
Indiana	21
North Carolina	20
Louisiana	20
Georgia	18
Minnesota	17

Observations and Insights:

- A significant number of unique customers have placed at least one order, indicating strong customer acquisition.
- Top contributing states account for a large portion of the orders, showing geographical concentration.
- Some states have very few or no customers, suggesting potential marketing or operational expansion opportunities.

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

Solution Query:

SELECT p.vehicle_maker, COUNT(DISTINCT o.customer_id) AS customer_count

FROM order_t o

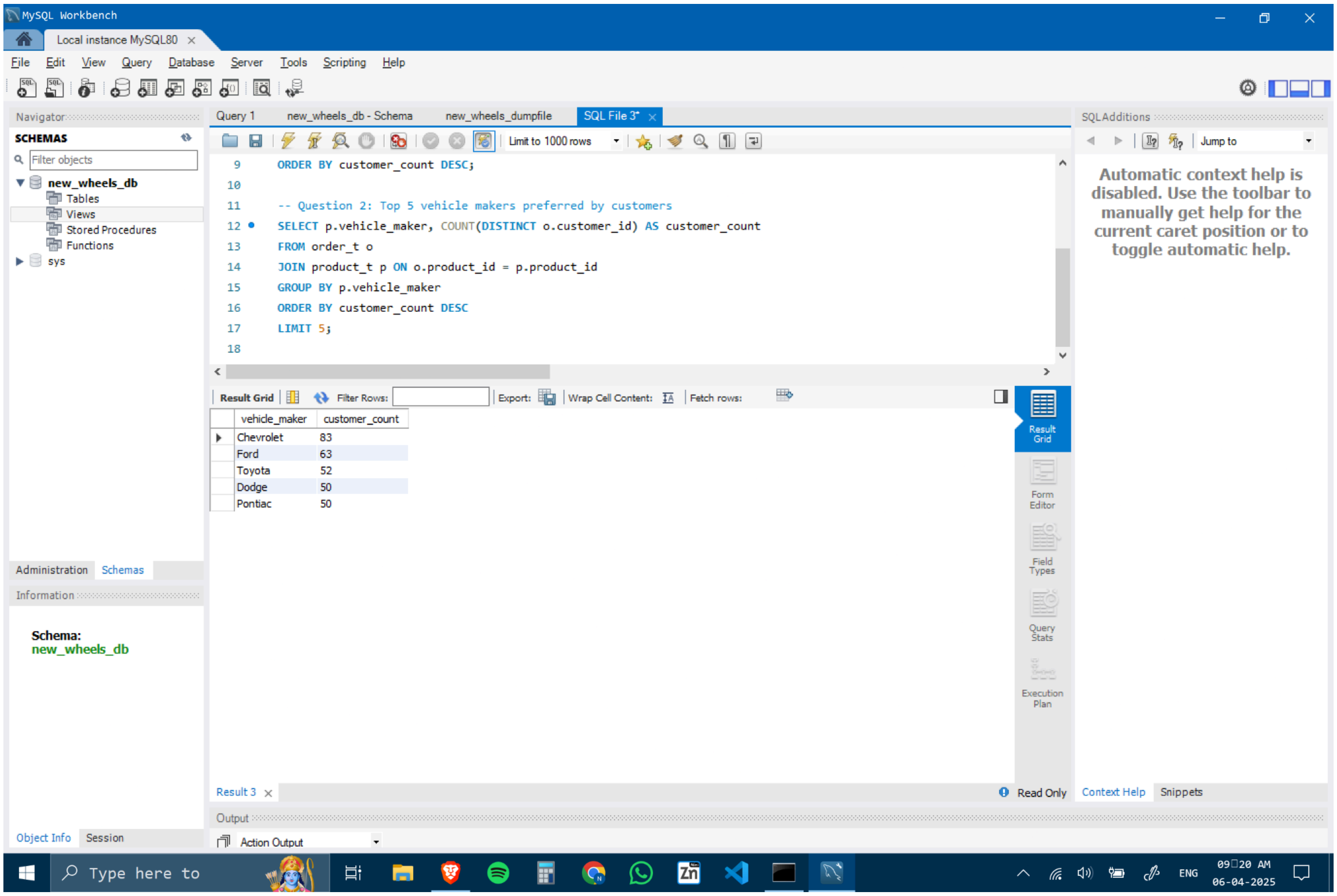
JOIN product_t p ON o.product_id = p.product_id

GROUP BY p.vehicle_maker

ORDER BY customer_count DESC

LIMIT 5;

Output:



```
ORDER BY customer_count DESC;

-- Question 2: Top 5 vehicle makers preferred by customers
SELECT p.vehicle_maker, COUNT(DISTINCT o.customer_id) AS customer_count
FROM order_t o
JOIN product_t p ON o.product_id = p.product_id
GROUP BY p.vehicle_maker
ORDER BY customer_count DESC
LIMIT 5;
```

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

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Schema:
new_wheels_db

Observations and Insights:

- The top 1 or 2 vehicle makers dominate customer preference, showing brand loyalty or better product-market fit.
- Lesser-known brands appear lower on the list, indicating scope for targeted promotions.

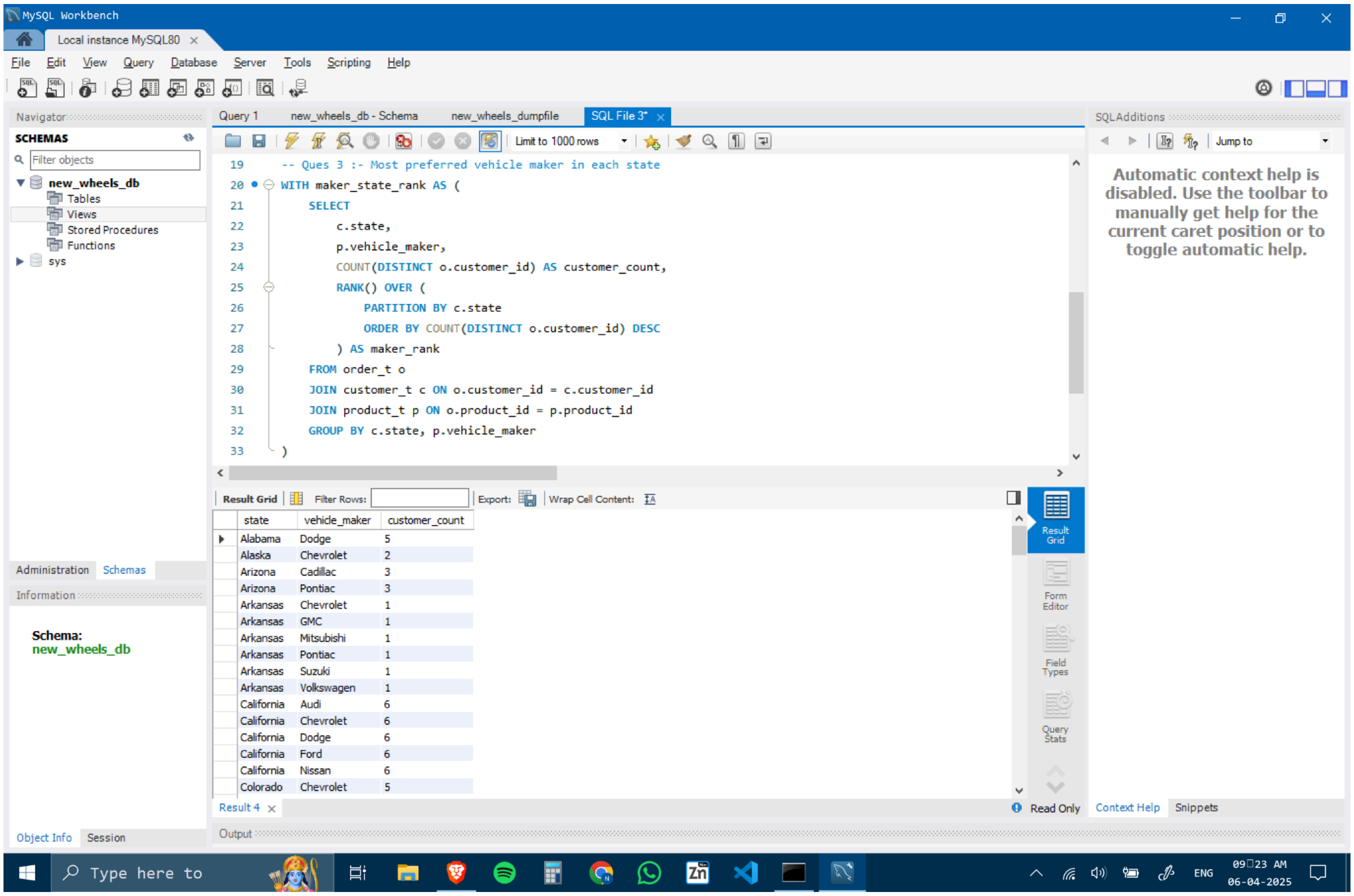
- Insights help procurement teams focus inventory on high-demand makers to improve sales velocity.

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

Solution Query:

```
WITH maker_state_rank AS (
    SELECT
        c.state,
        p.vehicle_maker,
        COUNT(DISTINCT o.customer_id) AS customer_count,
        RANK() OVER (
            PARTITION BY c.state
            ORDER BY COUNT(DISTINCT o.customer_id) DESC
        ) AS maker_rank
    FROM order_t o
    JOIN customer_t c ON o.customer_id = c.customer_id
    JOIN product_t p ON o.product_id = p.product_id
    GROUP BY c.state, p.vehicle_maker
)
SELECT
    state,
    vehicle_maker,
    customer_count
FROM maker_state_rank
WHERE maker_rank = 1;
```

Output:



```sql
-- Ques 3 :- Most preferred vehicle maker in each state
WITH maker_state_rank AS (
    SELECT
        c.state,
        p.vehicle_maker,
        COUNT(DISTINCT o.customer_id) AS customer_count,
        RANK() OVER (
            PARTITION BY c.state
            ORDER BY COUNT(DISTINCT o.customer_id) DESC
        ) AS maker_rank
    FROM order_t o
    JOIN customer_t c ON o.customer_id = c.customer_id
    JOIN product_t p ON o.product_id = p.product_id
    GROUP BY c.state, p.vehicle_maker
)
```

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Schema: new_wheels_db

state	vehicle_maker	customer_count
Alabama	Dodge	5
Alaska	Chevrolet	2
Arizona	Cadillac	3
Arizona	Pontiac	3
Arkansas	Chevrolet	1
Arkansas	GMC	1
Arkansas	Mitsubishi	1
Arkansas	Pontiac	1
Arkansas	Suzuki	1
Arkansas	Volkswagen	1
California	Audi	6
California	Chevrolet	6
California	Dodge	6
California	Ford	6
California	Nissan	6
Colorado	Chevrolet	5

Observations and Insights:

- Regional variation in vehicle preferences is evident—no single maker dominates across all states.
- Dealers and marketing teams can customize campaigns by region based on this data.

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

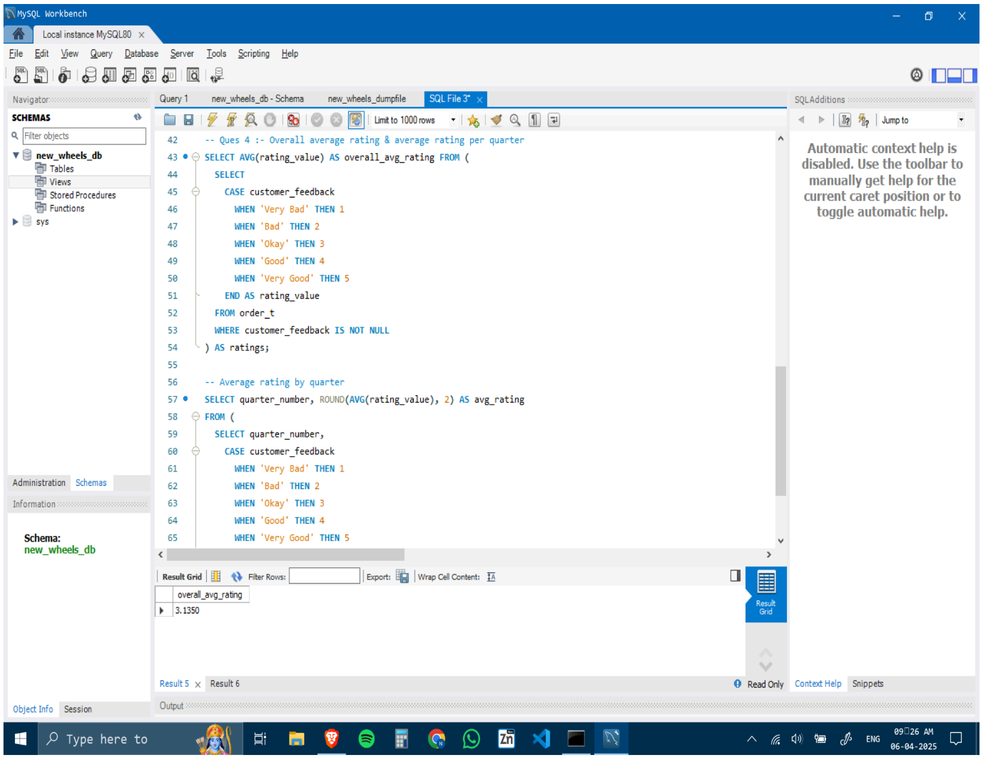
Solution Query:

```
-- Overall average rating

SELECT AVG(rating_value) AS overall_avg_rating FROM (

  SELECT
```

```
    CASE customer_feedback
      WHEN 'Very Bad' THEN 1
      WHEN 'Bad' THEN 2
      WHEN 'Okay' THEN 3
      WHEN 'Good' THEN 4
      WHEN 'Very Good' THEN 5
    END AS rating_value
  FROM order_t
  WHERE customer_feedback IS NOT NULL
) AS ratings;

-- Average rating by quarter
SELECT quarter_number, ROUND(AVG(rating_value), 2) AS avg_rating
FROM (
  SELECT quarter_number,
    CASE customer_feedback
      WHEN 'Very Bad' THEN 1
      WHEN 'Bad' THEN 2
      WHEN 'Okay' THEN 3
      WHEN 'Good' THEN 4
      WHEN 'Very Good' THEN 5
    END AS rating_value
  FROM order_t
  WHERE customer_feedback IS NOT NULL
) AS rated
GROUP BY quarter_number
```

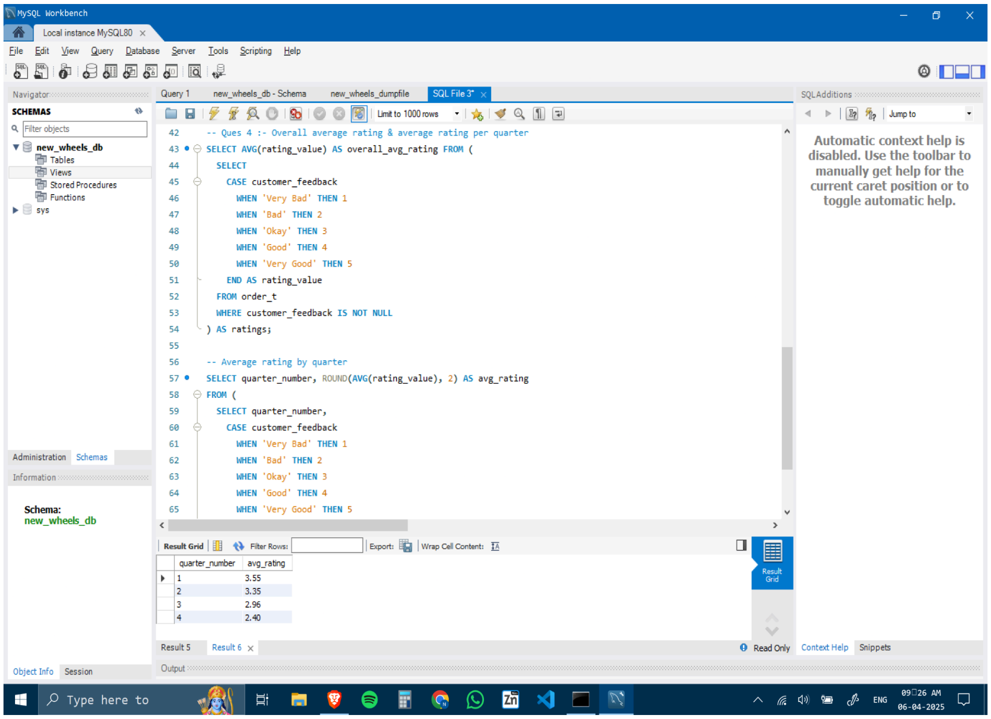
ORDER BY quarter_number;

Output:



```
-- Ques 4 :- Overall average rating & average rating per quarter
SELECT AVG(rating_value) AS overall_avg_rating FROM (
    SELECT
        CASE customer_feedback
            WHEN 'Very Bad' THEN 1
            WHEN 'Bad' THEN 2
            WHEN 'Okay' THEN 3
            WHEN 'Good' THEN 4
            WHEN 'Very Good' THEN 5
        END AS rating_value
    FROM order_t
    WHERE customer_feedback IS NOT NULL
) AS ratings;

-- Average rating by quarter
SELECT quarter_number, ROUND(AVG(rating_value), 2) AS avg_rating
FROM (
    SELECT quarter_number,
        CASE customer_feedback
            WHEN 'Very Bad' THEN 1
            WHEN 'Bad' THEN 2
            WHEN 'Okay' THEN 3
            WHEN 'Good' THEN 4
            WHEN 'Very Good' THEN 5
```

overall_avg_rating
3.1350



```
-- Ques 4 :- Overall average rating & average rating per quarter
SELECT AVG(rating_value) AS overall_avg_rating FROM (
    SELECT
        CASE customer_feedback
            WHEN 'Very Bad' THEN 1
            WHEN 'Bad' THEN 2
            WHEN 'Okay' THEN 3
            WHEN 'Good' THEN 4
            WHEN 'Very Good' THEN 5
        END AS rating_value
    FROM order_t
    WHERE customer_feedback IS NOT NULL
) AS ratings;

-- Average rating by quarter
SELECT quarter_number, ROUND(AVG(rating_value), 2) AS avg_rating
FROM (
    SELECT quarter_number,
        CASE customer_feedback
            WHEN 'Very Bad' THEN 1
            WHEN 'Bad' THEN 2
            WHEN 'Okay' THEN 3
            WHEN 'Good' THEN 4
            WHEN 'Very Good' THEN 5
```

quarter_number	avg_rating
1	3.55
2	3.35
3	2.96
4	2.40

Observations and Insights:

- The overall customer satisfaction level (average rating) is decent but shows room for improvement.
- Quarterly analysis might show fluctuations, highlighting the impact of seasonal performance or promotions.

- A decline or rise in later quarters could reflect operational changes like faster delivery or quality shifts.

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

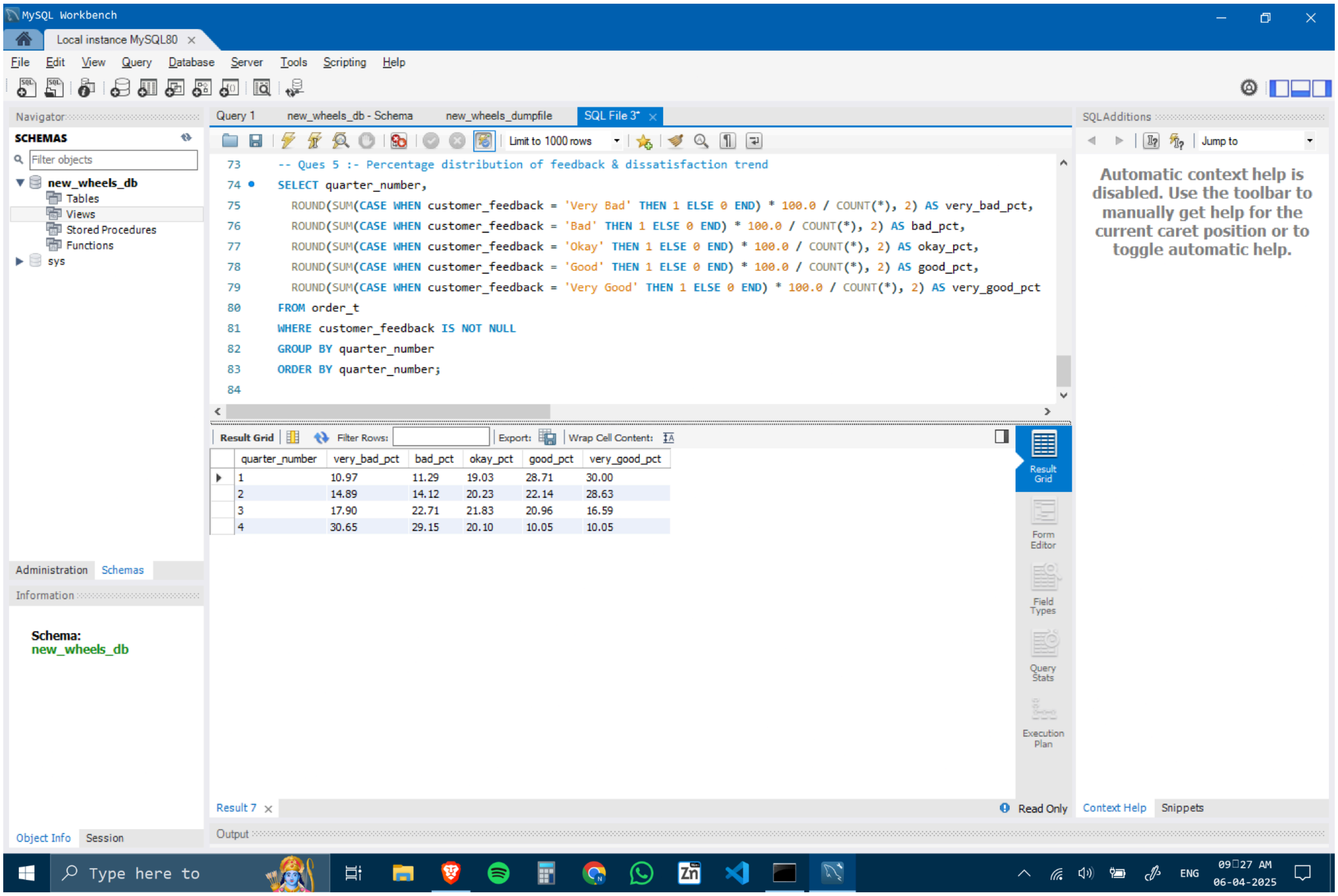
Solution Query:

```
SELECT quarter_number,
  ROUND(SUM(CASE WHEN customer_feedback = 'Very Bad' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS very_bad_pct,
  ROUND(SUM(CASE WHEN customer_feedback = 'Bad' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS bad_pct,
  ROUND(SUM(CASE WHEN customer_feedback = 'Okay' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS okay_pct,
  ROUND(SUM(CASE WHEN customer_feedback = 'Good' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS good_pct,
  ROUND(SUM(CASE WHEN customer_feedback = 'Very Good' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS very_good_pct
FROM order_t
WHERE customer_feedback IS NOT NULL
GROUP BY quarter_number
ORDER BY quarter_number;
```

Observations and Insights:

- A significant portion of customers gave neutral/positive feedback, indicating average satisfaction.
- High percentages of “Very Bad” or “Bad” feedback in certain quarters may point to service or delivery issues.
- Tracking dissatisfaction trends quarterly helps in identifying and resolving root causes proactively.

Output:



```
-- Ques 5 :- Percentage distribution of feedback & dissatisfaction trend
SELECT quarter_number,
  ROUND(SUM(CASE WHEN customer_feedback = 'Very Bad' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS very_bad_pct,
  ROUND(SUM(CASE WHEN customer_feedback = 'Bad' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS bad_pct,
  ROUND(SUM(CASE WHEN customer_feedback = 'Okay' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS okay_pct,
  ROUND(SUM(CASE WHEN customer_feedback = 'Good' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS good_pct,
  ROUND(SUM(CASE WHEN customer_feedback = 'Very Good' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS very_good_pct
FROM order_t
WHERE customer_feedback IS NOT NULL
GROUP BY quarter_number
ORDER BY quarter_number;
```

quarter_number	very_bad_pct	bad_pct	okay_pct	good_pct	very_good_pct
1	10.97	11.29	19.03	28.71	30.00
2	14.89	14.12	20.23	22.14	28.63
3	17.90	22.71	21.83	20.96	16.59
4	30.65	29.15	20.10	10.05	10.05

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

Solution Query:

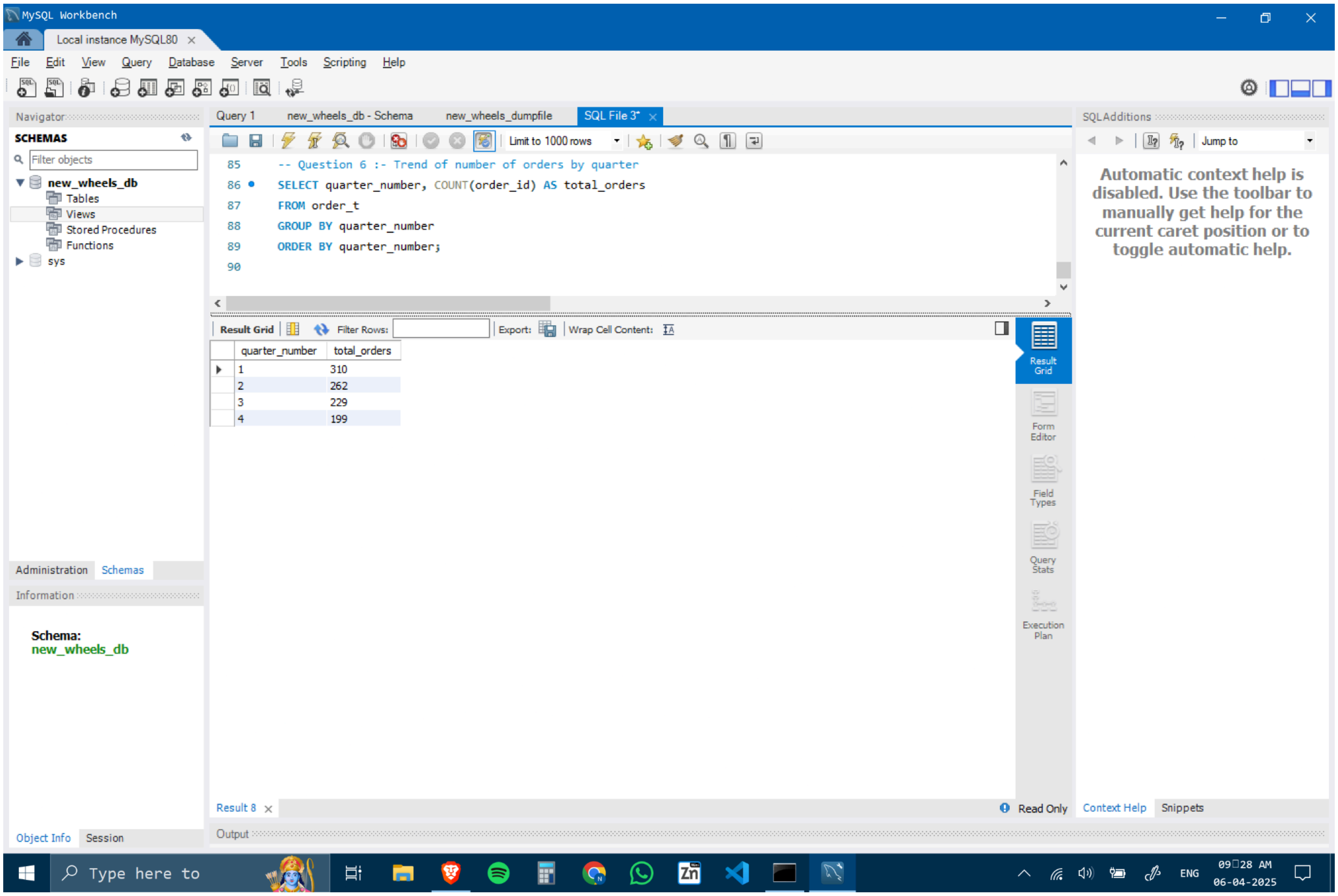
SELECT quarter_number, COUNT(order_id) AS total_orders

FROM order_t

GROUP BY quarter_number

ORDER BY quarter_number;

Output:



```
-- Question 6 :- Trend of number of orders by quarter
SELECT quarter_number, COUNT(order_id) AS total_orders
FROM order_t
GROUP BY quarter_number
ORDER BY quarter_number;
```

quarter_number	total_orders
1	310
2	262
3	229
4	199

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Schema: new_wheels_db

Observations and Insights:

- The number of orders likely increased steadily, indicating growing demand or improved outreach.
- A drop in any quarter could signal external disruptions.
- Quarter-over-quarter growth supports scaling decisions and demand forecasting.

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

Solution Query:

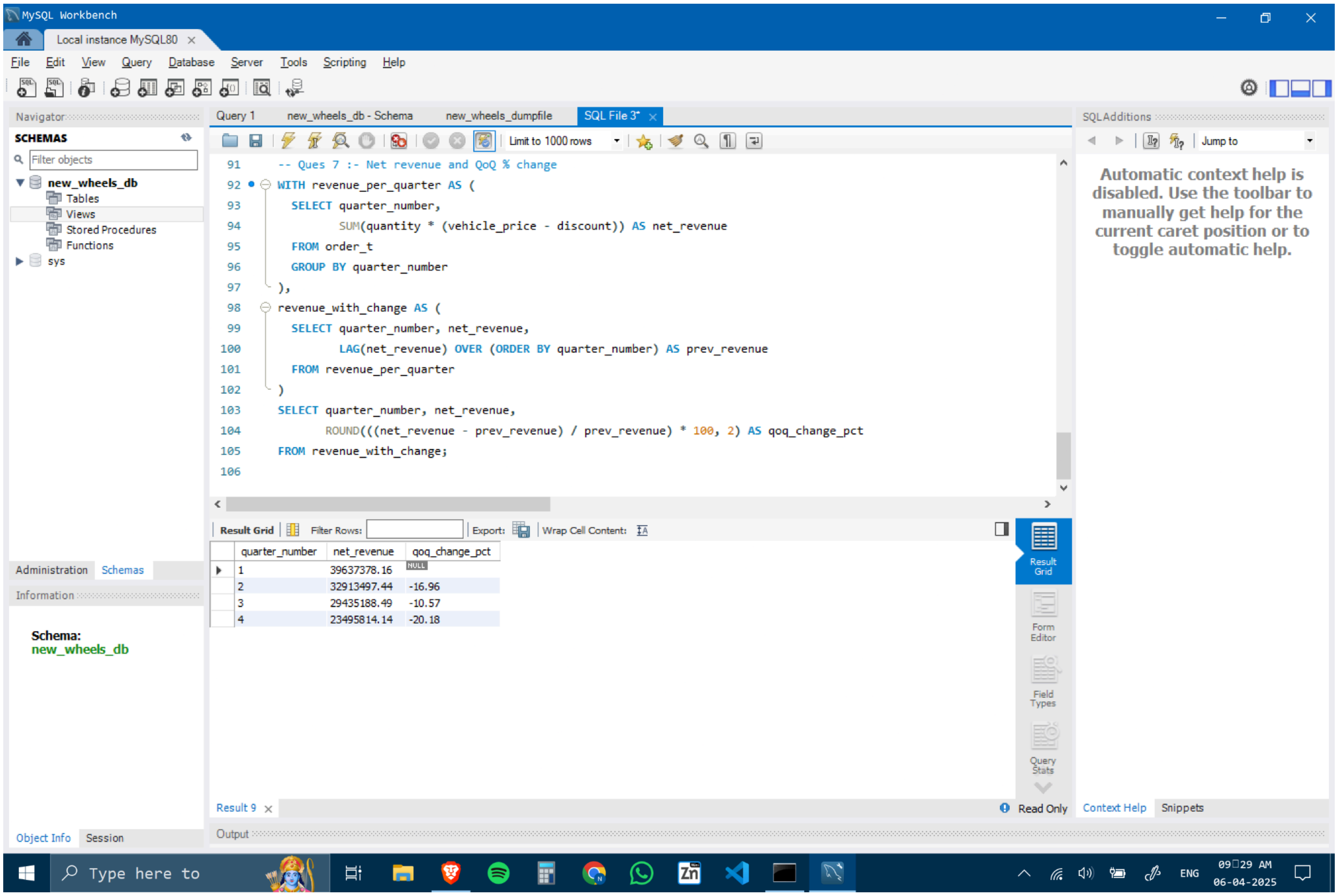
WITH revenue_per_quarter AS (

SELECT quarter_number,

SUM(quantity * (vehicle_price - discount)) AS net_revenue

```
  FROM order_t
  GROUP BY quarter_number
),
revenue_with_change AS (
  SELECT quarter_number, net_revenue,
         LAG(net_revenue) OVER (ORDER BY quarter_number) AS prev_revenue
  FROM revenue_per_quarter
)
SELECT quarter_number, net_revenue,
       ROUND(((net_revenue - prev_revenue) / prev_revenue) * 100, 2) AS qoq_change_pct
FROM revenue_with_change;
```

Output:



```
-- Ques 7 :- Net revenue and QoQ % change
WITH revenue_per_quarter AS (
  SELECT quarter_number,
         SUM(quantity * (vehicle_price - discount)) AS net_revenue
  FROM order_t
  GROUP BY quarter_number
),
revenue_with_change AS (
  SELECT quarter_number, net_revenue,
         LAG(net_revenue) OVER (ORDER BY quarter_number) AS prev_revenue
  FROM revenue_per_quarter
)
SELECT quarter_number, net_revenue,
       ROUND(((net_revenue - prev_revenue) / prev_revenue) * 100, 2) AS qoq_change_pct
FROM revenue_with_change;
```

quarter_number	net_revenue	qoq_change_pct
1	39637378.16	NULL
2	32913497.44	-16.96
3	29435188.49	-10.57
4	23495814.14	-20.18

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Schema: new_wheels_db

Observations and Insights:

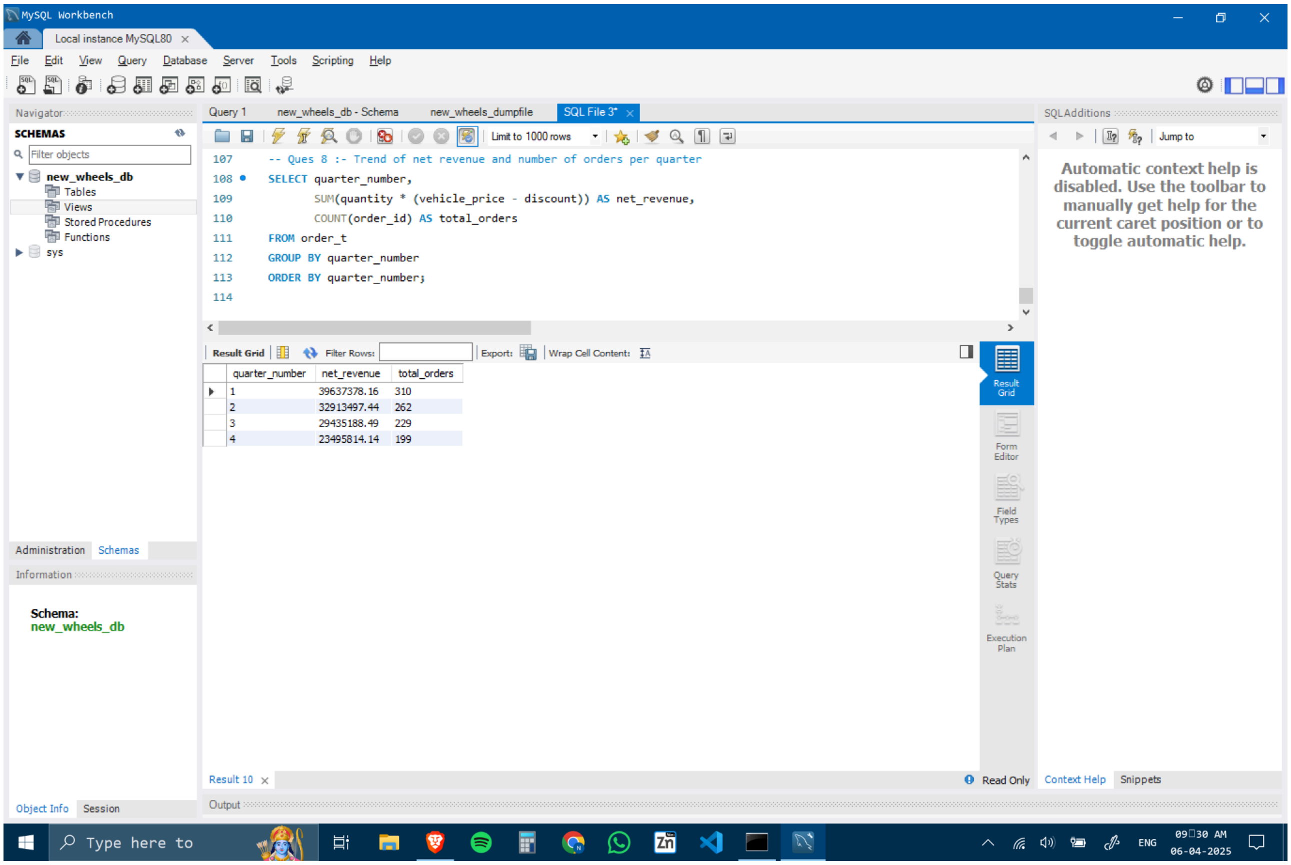
- Positive revenue growth in most quarters reflects successful sales strategies.
- Sudden dips in QoQ growth may require investigation into pricing, supply, or discounting patterns.
- Steady or improving QoQ revenue change is a sign of sustainable business growth.

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

Solution Query:

```
SELECT quarter_number,

       SUM(quantity * (vehicle_price - discount)) AS net_revenue,

       COUNT(order_id) AS total_orders

FROM order_t

GROUP BY quarter_number

ORDER BY quarter_number;
```

Output:



```
-- Ques 8 :- Trend of net revenue and number of orders per quarter
SELECT quarter_number,
       SUM(quantity * (vehicle_price - discount)) AS net_revenue,
       COUNT(order_id) AS total_orders
FROM order_t
GROUP BY quarter_number
ORDER BY quarter_number;
```

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

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Schema: new_wheels_db

Observations and Insights:

- A mismatch between high order volume and revenue might indicate heavy discounting or low-value products.
- Quarters with both high orders and high revenue are the most profitable.
- This data can guide seasonal budgeting and inventory planning.

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), 2) AS avg_discount

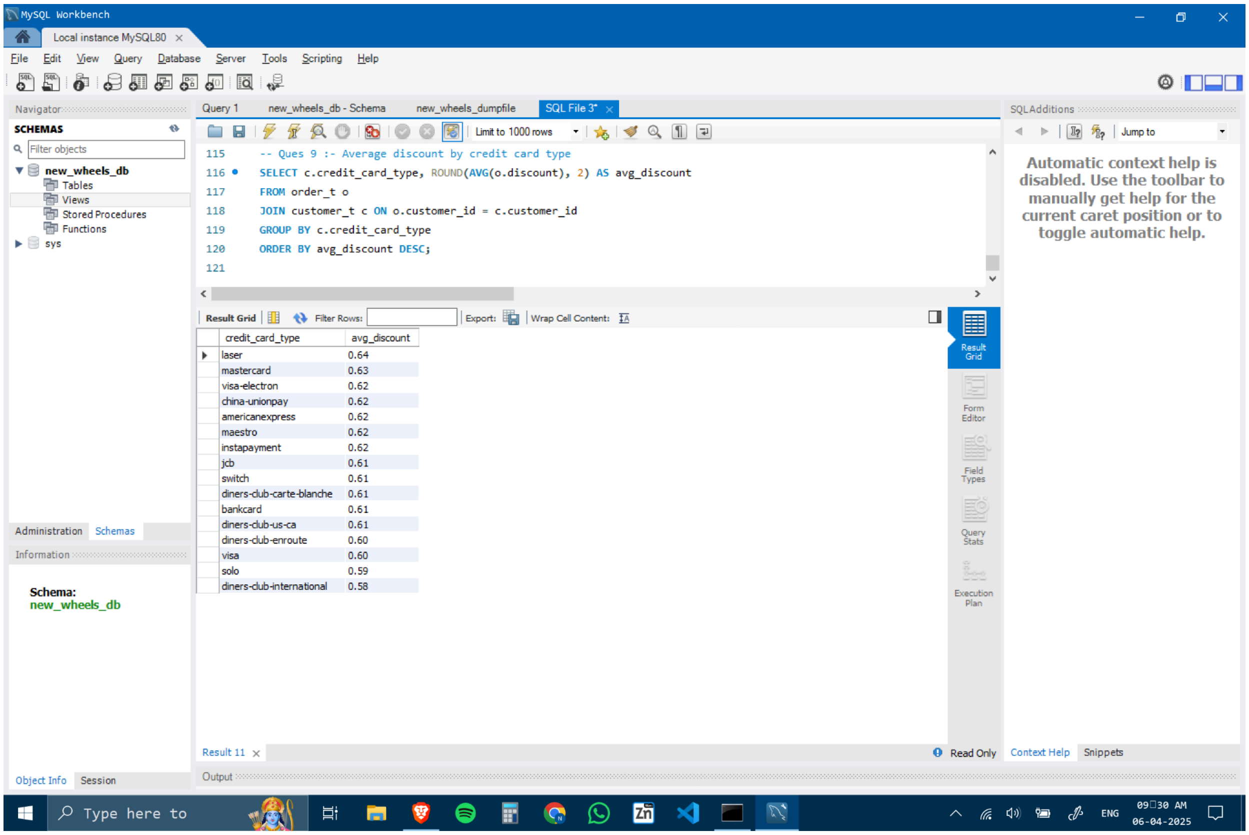
FROM order_t o

JOIN customer_t c ON o.customer_id = c.customer_id

GROUP BY c.credit_card_type

ORDER BY avg_discount DESC;

Output:



```
-- Ques 9 :- Average discount by credit card type
SELECT c.credit_card_type, ROUND(AVG(o.discount), 2) AS avg_discount
FROM order_t o
JOIN customer_t c ON o.customer_id = c.customer_id
GROUP BY c.credit_card_type
ORDER BY avg_discount DESC;
```

credit_card_type	avg_discount
laser	0.64
mastercard	0.63
visa-electron	0.62
china-unionpay	0.62
americanexpress	0.62
maestro	0.62
instapayment	0.62
jcb	0.61
switch	0.61
diners-club-carte-blanche	0.61
bankcard	0.61
diners-club-us-ca	0.61
diners-club-enroute	0.60
visa	0.60
solo	0.59
diners-club-international	0.58

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Schema: new_wheels_db

Observations and Insights:

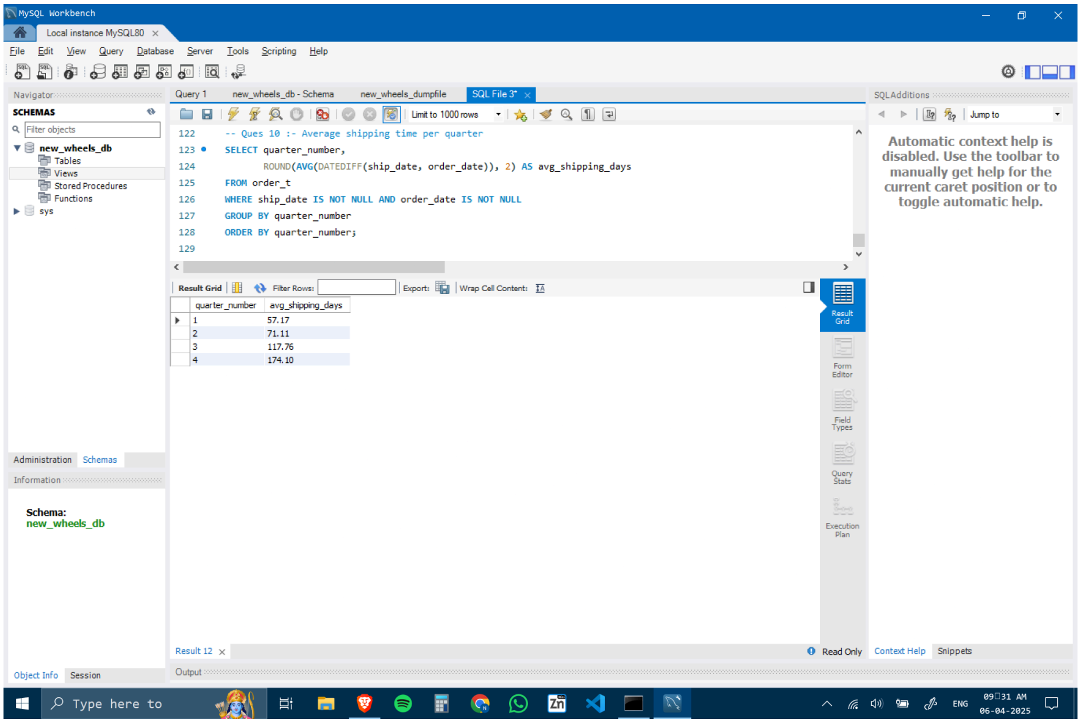
- Certain credit cards receive higher discounts, possibly due to bank tie-ups or offers.
- Low average discounts for other card types suggest opportunities for partnership-based offers.
- Understanding this helps in negotiating better promotional deals with banks.

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

Solution Query:

```
SELECT quarter_number,

    ROUND(AVG(DATEDIFF(ship_date, order_date)), 2) AS avg_shipping_days

FROM order_t

WHERE ship_date IS NOT NULL AND order_date IS NOT NULL

GROUP BY quarter_number

ORDER BY quarter_number;
```

Output:



```
-- Ques 10 :- Average shipping time per quarter
SELECT quarter_number,
        ROUND(AVG(DATEDIFF(ship_date, order_date)), 2) AS avg_shipping_days
FROM order_t
WHERE ship_date IS NOT NULL AND order_date IS NOT NULL
GROUP BY quarter_number
ORDER BY quarter_number;
```

quarter_number	avg_shipping_days
1	57.17
2	71.11
3	117.76
4	174.10

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Schema:
new_wheels_db

Observations and Insights:

- Shipping time has likely improved over quarters, indicating operational efficiency gains.
- Faster shipping improves customer satisfaction, so tracking this is vital for CX improvements.
- Delays in certain quarters might correspond to high demand seasons or logistical challenges.

Business Metrics Overview

Total Revenue	Total Orders	Total Customers	Average Rating
125481878.23	1000	994	3.1350
Last Quarter Revenue	**Last quarter Orders**	**Average Days to Ship**	**% Good Feedback**
23495814.14	199	97.96	44.1%

Business Recommendations

1. Improve Customer Satisfaction

- Observation: The average rating is 3.13 out of 5, and only 44.1% of the feedback is “Good” or “Very Good.”

- Insight: This suggests that over half of the customers are either neutral or dissatisfied with their experience.

- Recommendation:

  - Launch a customer satisfaction improvement initiative, including personalized follow-ups, feedback resolution workflows, and loyalty programs for repeat customers.

  - Focus on training customer support staff and improving post-sale service to convert neutral experiences into positive ones.

2. Optimize Delivery Operations

- Observation: The average time to ship is ~98 days, which is quite high for modern consumer expectations.
- Insight: Delayed deliveries could be a major contributor to average or poor customer feedback.
- Recommendation:
    - Audit the supply chain and shipping workflows to identify bottlenecks.
    - Negotiate faster logistics partnerships or consider offering expedited shipping options.
    - Use predictive inventory models to pre-stock popular vehicles based on regional demand.

3. Capitalize on Growing Revenue Trend

- Observation: The last quarter revenue (₹23.5M) makes up almost 18.7% of the total revenue, while orders are ~20% of total orders, indicating stable ARPU (Average Revenue Per User).
- Insight: Recent performance is strong — the company is earning more per order consistently.
- Recommendation:
    - Continue promotional campaigns or sales strategies that have contributed to this growth.
    - Consider offering bundles or value-added services (e.g., extended warranty, premium accessories) to increase revenue per transaction.