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
    a.state,
    a.no_of_customers,
    b.total_no_of_customers_placed_orders
FROM
    (SELECT
        state,
        COUNT(DISTINCT customer_id) AS no_of_customers
     FROM
        gl_project.customer_t
     GROUP BY
        state) a
INNER JOIN
    (SELECT
        COUNT(DISTINCT customer_id) AS total_no_of_customers_placed_orders
     FROM
        gl_project.order_t) b
ON
    1 = 1
ORDER BY
    a.no_of_customers DESC;
```

Output:

state	no_of_customers	total_no_of_customers_placed_orders_across_all_states
California	97	994
Texas	97	994
Florida	86	994
New York	69	994
District of Columbia	35	994
Colorado	33	994
Ohio	33	994
Alabama	29	994
Washington	28	994
Arizona	26	994
Pennsylvania	25	994
Illinois	25	994
Virginia	24	994
Tennessee	23	994
Missouri	23	994
Connecticut	22	994
Indiana	21	994
North Carolina	20	994
Louisiana	20	994
Georgia	18	994
Michigan	17	994
Nevada	17	994
Minnesota	17	994

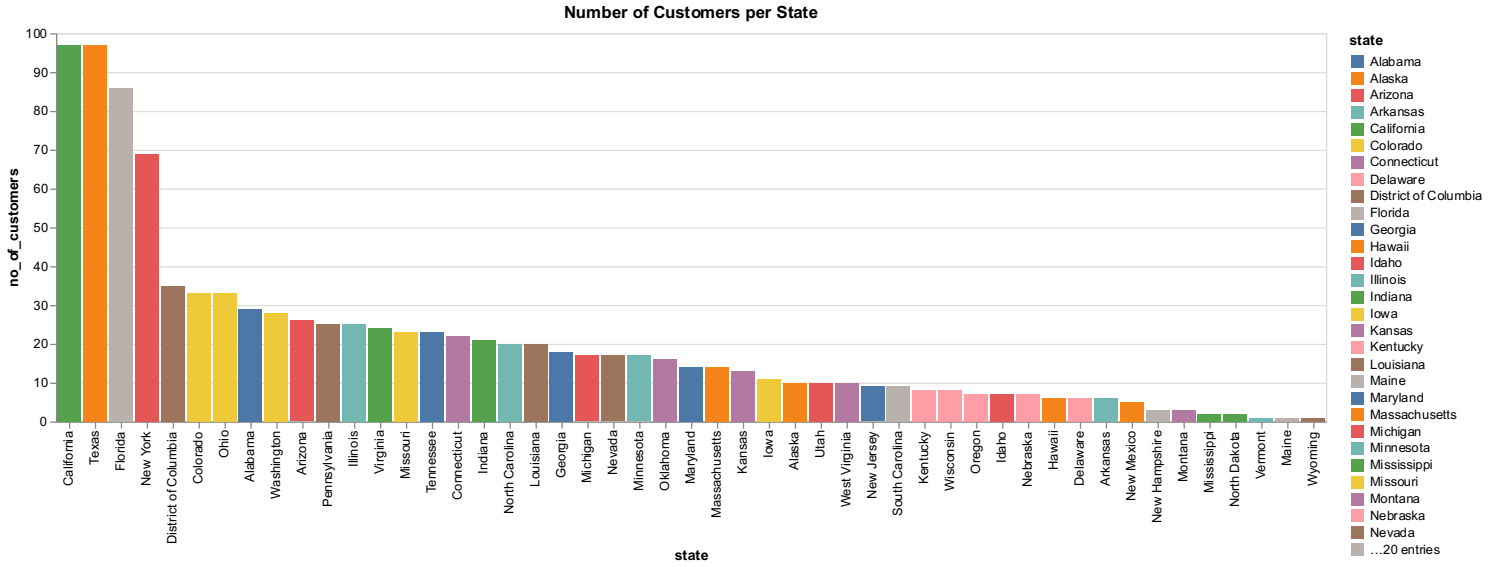
No of States:

```
SELECT
    COUNT(DISTINCT state) AS no_of_states
FROM
    gl_project.customer_t;
```

Output:

no_of_states
49

Graph:



Observations and Insights:

- There are around 994 customers in total who placed order
- These customers are distributed across 49 states
- **California** and **Texas** have highest number customers 97 each followed by Florida at 86.
- Top 4 states contain most no of customer. **(California, Texas, Florida, New York)**

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

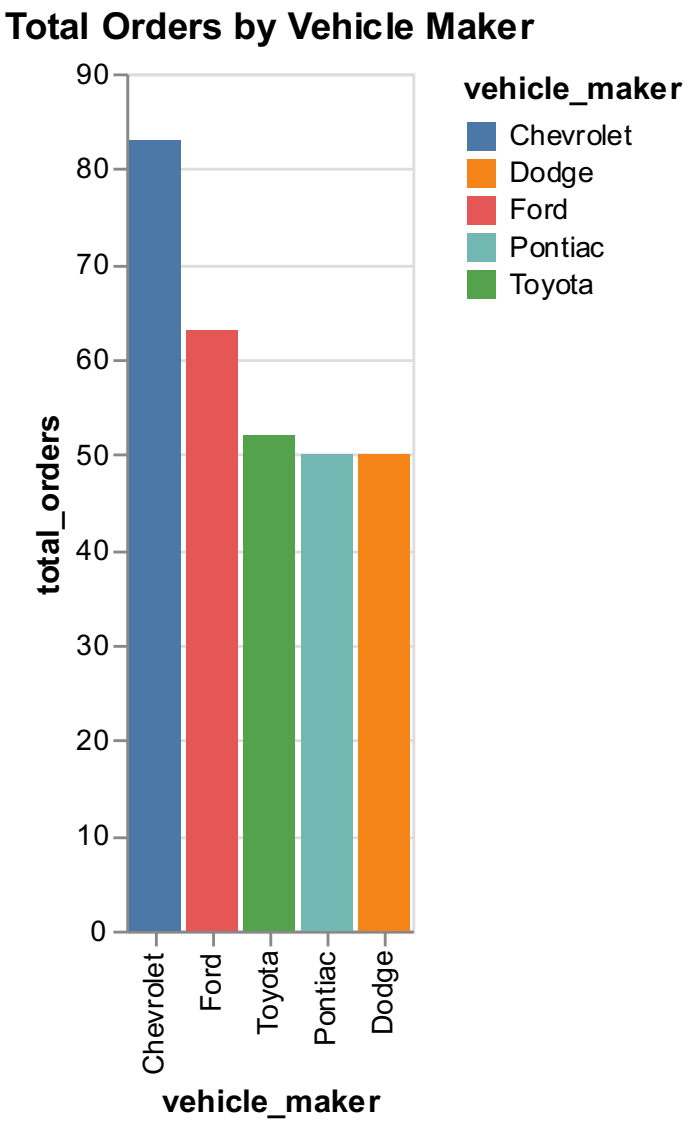
Solution Query:

```
SELECT
    P1.vehicle_maker,
    COUNT(O.order_id) AS total_orders
FROM
    gl_project.product_t P1
INNER JOIN
    gl_project.order_t O
ON
    P1.product_id = O.product_id
GROUP BY
    P1.vehicle_maker
ORDER BY
    total_orders DESC
LIMIT 5;
```

Output:

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

Graph:



Observations and Insights:

- Chevrolet leads as the most preferred vehicle maker with the highest customer count, followed by Ford.

- Toyota, Pontiac, and Dodge have an approximately similar number of customers, indicating comparable popularity.
- Try to cash in the demand for Chevrolet and Ford by maintaining a sufficient inventory and offering competitive price.
- Develop targeted approaches to elevate customer engagement and satisfaction for Toyota, Pontiac, and Dodge to boost their market share and strengthen their positions

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

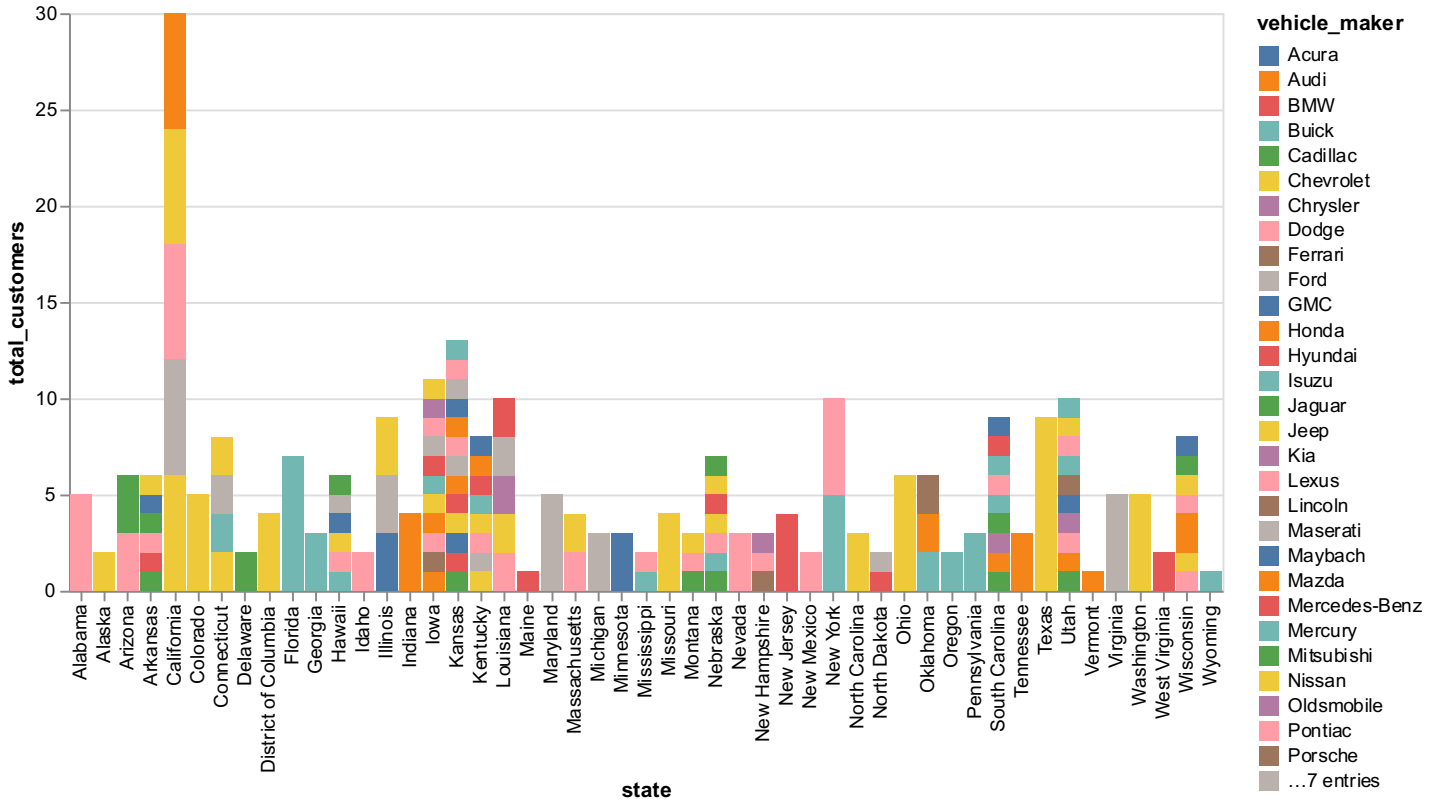
Solution Query:

```
WITH StateVehicleOrders AS (
    SELECT
        C.state,
        P.vehicle_maker,
        COUNT(C.customer_id) AS total_customers
    FROM
        gl_project.product_t P
    INNER JOIN
        gl_project.order_t O
    ON
        P.product_id = O.product_id
    INNER JOIN
        gl_project.customer_t C
    ON
        O.customer_id = C.customer_id
    GROUP BY
        C.state, P.vehicle_maker
),
RankedStateVehicles AS (
    SELECT
        state,
        vehicle_maker,
        total_customers,
        RANK() OVER (PARTITION BY state ORDER BY total_customers
DESC) AS vehicle_rank
    FROM
        StateVehicleOrders
)
SELECT
    state,
    vehicle_maker,
    total_customers
FROM
    RankedStateVehicles
WHERE
    vehicle_rank = 1
ORDER BY
    total_customers DESC;
```

Output:

state	vehicle_maker	total_customers
Texas	Chevrolet	9
Florida	Toyota	7
California	Chevrolet	6
California	Nissan	6
California	Dodge	6
California	Ford	6
California	Audi	6
Ohio	Chevrolet	6
Alabama	Dodge	5
Colorado	Chevrolet	5
Maryland	Ford	5

Graph:



Observations and Insights:

- The preferred vehicle in most of the state is Chevrolet.
- California has highest customer.
- Chevrolet as a brand is most proffered in Texas and California.

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

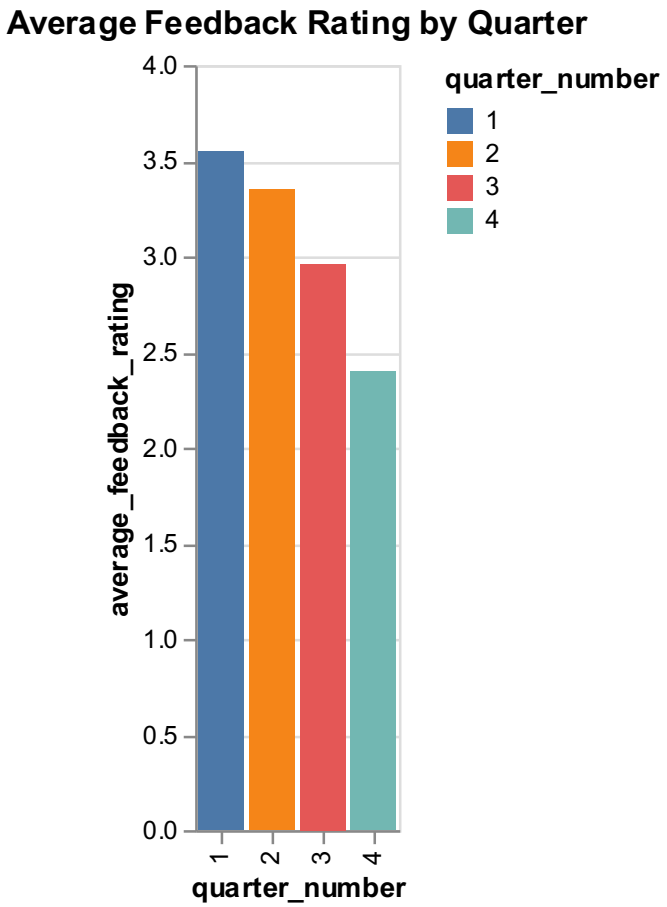
Solution Query:

```
WITH FeedbackRatings AS (
    SELECT
        customer_feedback,
        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 feedback_rating
    FROM
        gl_project.order_t
),
QuarterlyAverage AS (
    SELECT
        quarter_number,
        ROUND(AVG(feedback_rating), 2) AS average_feedback_rating
    FROM
        FeedbackRatings
    GROUP BY
        quarter_number
),
OverallAverage AS (
    SELECT
        ROUND(AVG(feedback_rating), 2) AS overall_feedback_rating
    FROM
        FeedbackRatings
)
SELECT
    qa.quarter_number,
    qa.average_feedback_rating,
    oa.overall_feedback_rating
FROM
    QuarterlyAverage qa
INNER JOIN
    OverallAverage oa
ON
    1=1
ORDER BY
    qa.quarter_number ASC;
```

Output:

quarter_number	average_feedback_rating	overall_feedback_rating
1	3.55	3.14
2	3.35	3.14
3	2.96	3.14
4	2.40	3.14

Graph:



Observations and Insights:

- Quarter 1 has the highest average feedback rating, slightly above **3.5**. Quarter 2 follows closely, with an average rating like Quarter 1. Quarter 3 shows a moderate decline, with an average rating just above **3.0**. Quarter 4 has the lowest average rating, below **3.0**, indicating a noticeable drop in feedback performance.
- Around -1.15 rating decreased from Q1 to Q4
- There is a clear downward trend in average feedback ratings as the year progresses.
- Overall feedback ratting across all is around 3.14
- New wheels need to investigate the reasons for the decline in ratings in Quarters 3 and 4 to improve feedback scores.

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

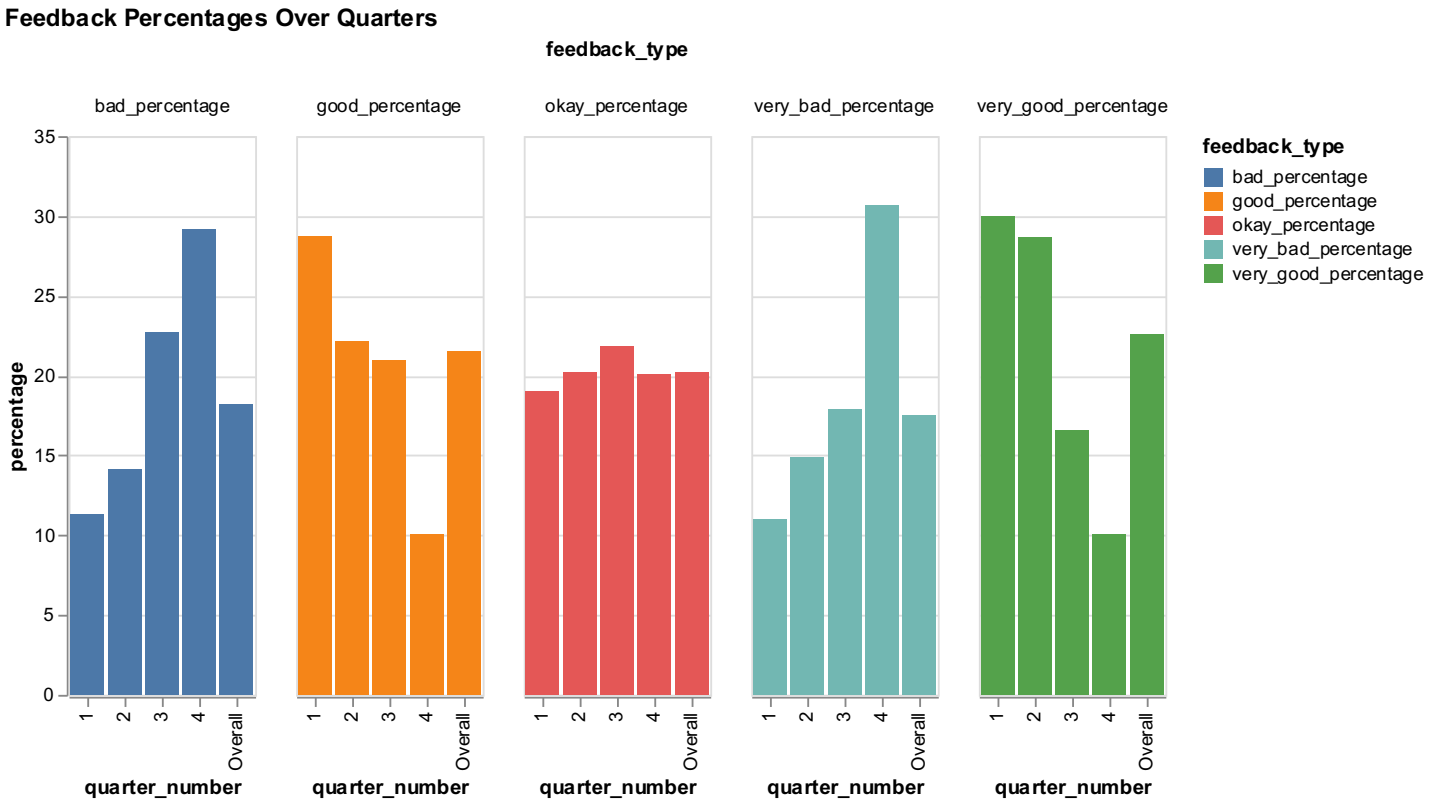
Solution Query:

```
WITH feedback AS (
    SELECT
        COALESCE(quarter_number, 'Overall') AS quarter_number,
        COUNT(customer_feedback) AS total_feedback,
        ROUND(SUM(CASE WHEN customer_feedback = 'very bad' THEN 1 ELSE 0
END) * 100.0 / COUNT(customer_feedback), 2) AS very_bad_percentage,
        ROUND(SUM(CASE WHEN customer_feedback = 'bad' THEN 1 ELSE 0 END) *
100.0 / COUNT(customer_feedback), 2) AS bad_percentage,
        ROUND(SUM(CASE WHEN customer_feedback = 'okay' THEN 1 ELSE 0 END) *
100.0 / COUNT(customer_feedback), 2) AS okay_percentage,
        ROUND(SUM(CASE WHEN customer_feedback = 'good' THEN 1 ELSE 0 END) *
100.0 / COUNT(customer_feedback), 2) AS good_percentage,
        ROUND(SUM(CASE WHEN customer_feedback = 'very good' THEN 1 ELSE 0
END) * 100.0 / COUNT(customer_feedback), 2) AS very_good_percentage
    FROM
        gl_project.order_t
    GROUP BY ROLLUP(quarter_number)
)
SELECT *
FROM feedback
ORDER BY
    quarter_number;
```

Output:

quarter_number	total_feedback	very_bad_percentage	bad_percenta...	okay_percentage	good_percentage	very_good_percentage
1	310	10.97	11.29	19.03	28.71	30.00
2	262	14.89	14.12	20.23	22.14	28.63
3	229	17.90	22.71	21.83	20.96	16.59
4	199	30.65	29.15	20.10	10.05	10.05
Overall	1000	17.50	18.20	20.20	21.50	22.60

Graph:



Observations and Insights:

- **Feedback Decline**: Total feedback drops from 310 (Q1) to 199 (Q4).
- **Good Feedback Same for Q**1:We can see good feedback and very good feedback is around the same level in Q1 and bad feedback and very bad feedback around the same level in Q4.
- **Rising Dissatisfaction**: "Very Bad" feedback rises from **10.97%** (Q1) to **30.65%** (Q4), and "Bad" from **11.29%** to **29.15%**.
- **Falling Satisfaction**: "Very Good" drops from **30.00%** (Q1) to **10.05%** (Q4).
- **Trend**: Positive feedback decreases, while negative feedback increases over time.
- We can clearly see though the total no of feedback decreases but there is significant rise of negative feedback .
- **Action Needed**: Address dissatisfaction and enhance service quality in later quarters.

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

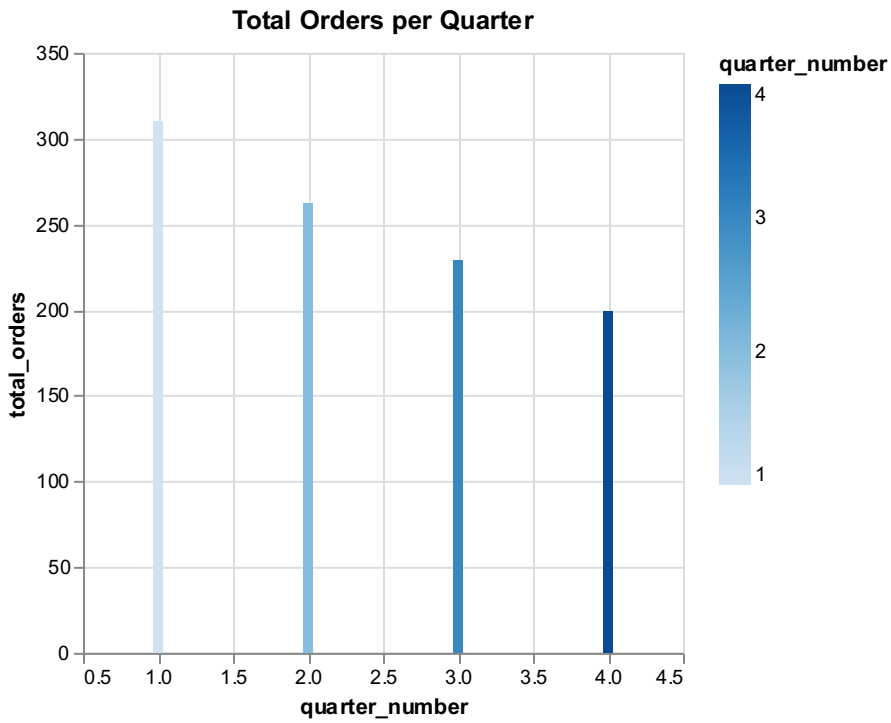
Solution Query:

```
SELECT
    quarter_number,
    count(order_id) as total_orders
FROM
    gl_project.order_t
GROUP BY
    quarter_number
ORDER BY
    quarter_number ASC;
```

Output:

quarter_numb...	total_orders
1	310
2	262
3	229
4	199

Graph:



Observations and Insights:

- Total orders declined across the four quarters, with Q1 being the highest and Q4 the lowest.
- This could be due to declining ratings from customer over the time.

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

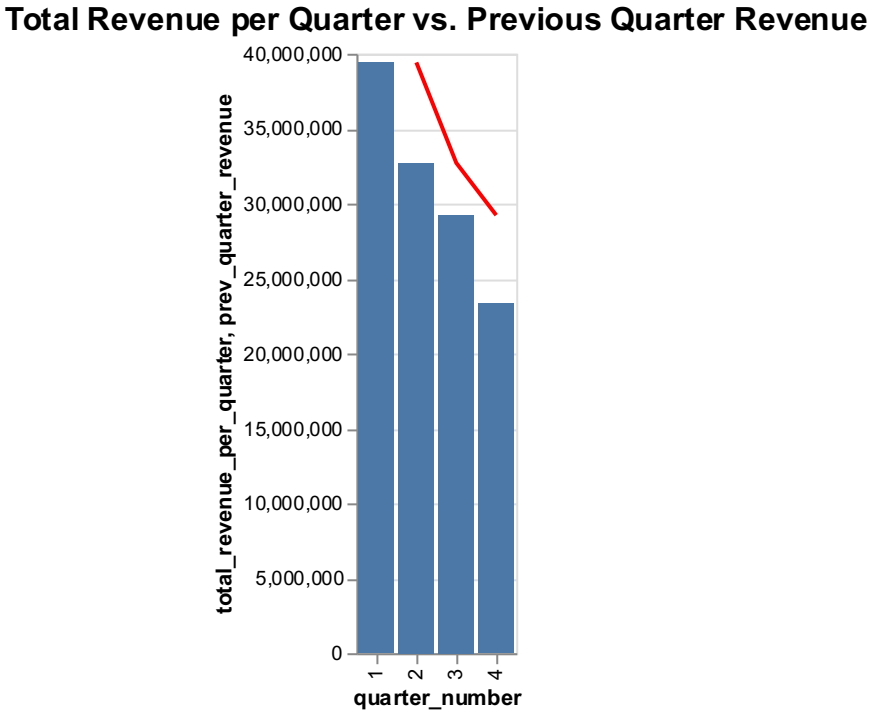
Solution Query:

```
WITH quarterly_revenue AS (
    SELECT
        quarter_number,
        ROUND(SUM(quantity * (vehicle_price -
((discount/100)*vehicle_price))), 0) AS total_revenue_per_quarter
    FROM
        gl_project.order_t
    GROUP BY
        quarter_number
    ORDER BY
        quarter_number ASC
),
revenue_with_lag AS (
    SELECT
        quarter_number,
        total_revenue_per_quarter,
        LAG(total_revenue_per_quarter, 1) OVER (ORDER BY quarter_number
ASC) AS prev_quarter_revenue
    FROM
        quarterly_revenue
),
overall_net_revenue AS (
    SELECT
        ROUND(SUM(quantity * (vehicle_price -
((discount/100)*vehicle_price))), 0) AS overall_net_revenue
    FROM
        gl_project.order_t
)
SELECT
    rw.quarter_number,
    rw.total_revenue_per_quarter,
    rw.prev_quarter_revenue,
    CASE
        WHEN rw.prev_quarter_revenue IS NOT NULL THEN
            (rw.total_revenue_per_quarter - rw.prev_quarter_revenue) /
rw.prev_quarter_revenue * 100
        ELSE
            0.0
    END AS quarter_percentage_change,
    onr.overall_net_revenue
FROM
    revenue_with_lag rw
    INNER JOIN overall_net_revenue onr ON 1=1
ORDER BY
    rw.quarter_number ASC;
```

Output:

quarter_number	total_revenue_per_quar...	prev_quarter_revenue	quarter_percentage_change	overall_net_revenue
1	39421580	NULL	0.0	124714086
2	32715830	39421580	-17.0104	124714086
3	29229896	32715830	-10.6552	124714086
4	23346780	29229896	-20.1271	124714086

Graph:



Observations and Insights:

- There is a consistent **decline in revenue** each quarter
- The most significant percentage decline occurred between Quarter 3 and Quarter 4 (**20.13%**) followed by Q1 to Q2
- In Q4 the revenue fell by 6 million which is around 20%

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

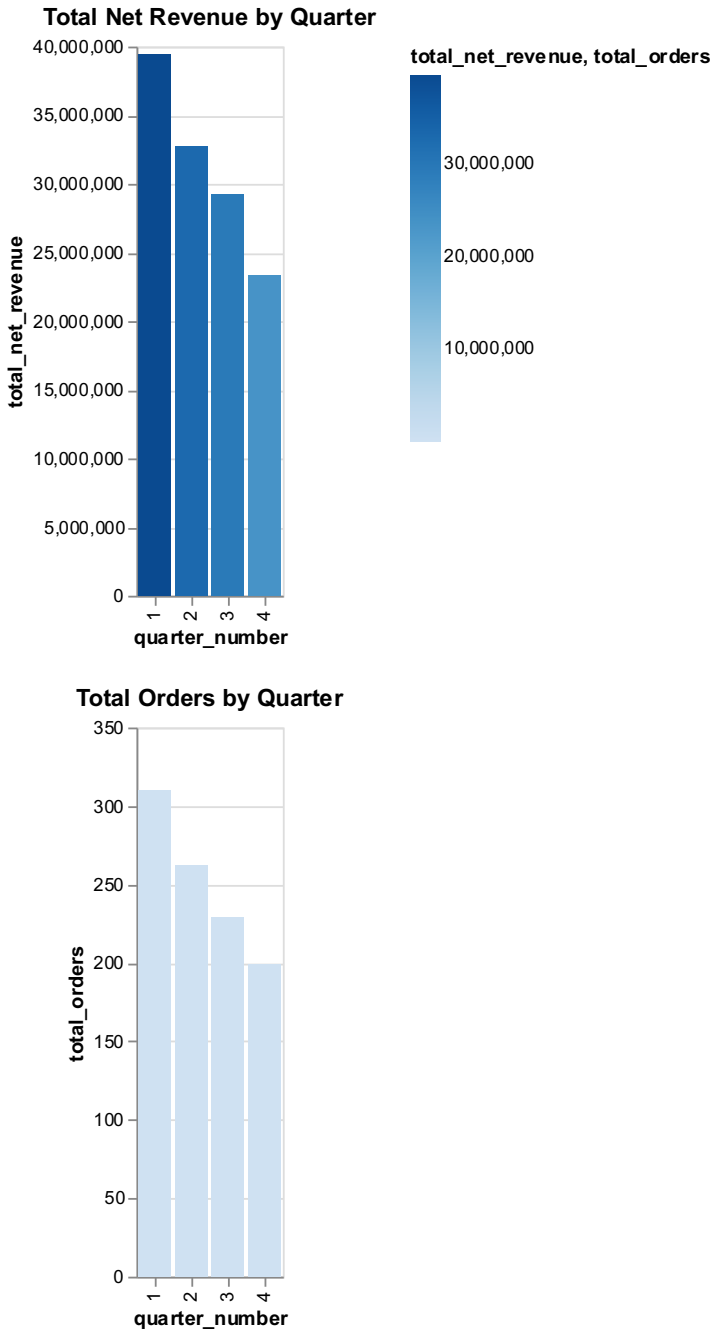
Solution Query:

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

Output:

quarter_numb...	total_net_reven...	total_orders
1	39421580	310
2	32715830	262
3	29229896	229
4	23346780	199

Graph:



Observations and Insights:

- **Trend**: Revenue and orders decrease from Q1 to Q4.
- While **no of orders** decreased from 310 to 199.The revenue reduced from $39 million approximately to $23 million approximately.

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

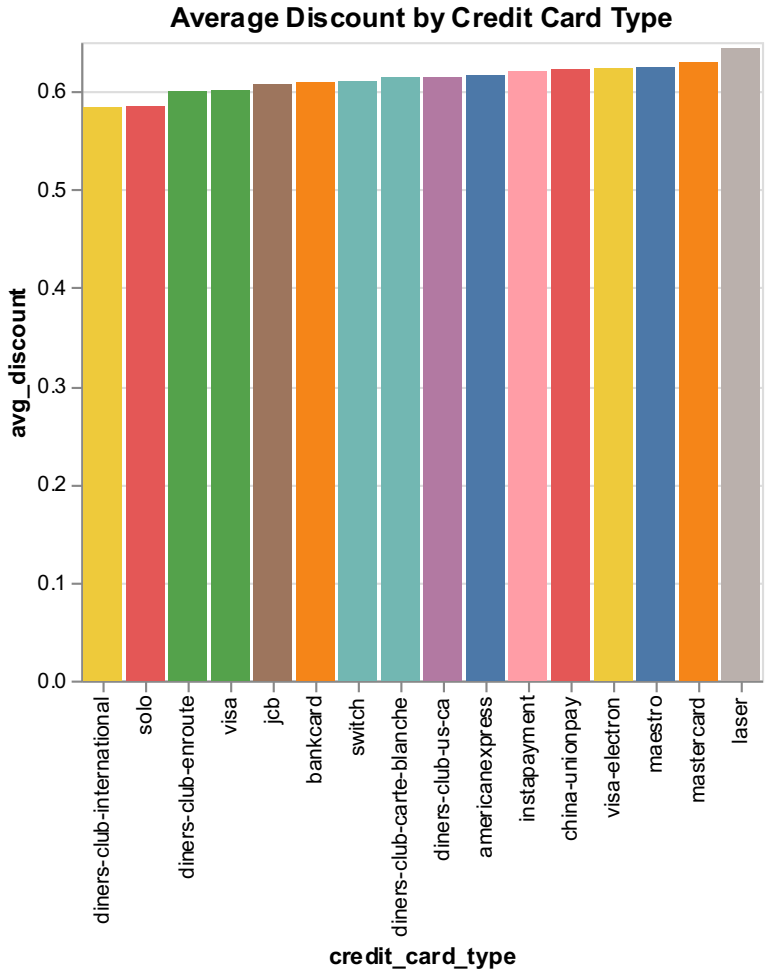
Solution Query:

```
SELECT
    C.credit_card_type,
    AVG(O.discount) AS avg_discount
FROM
    gl_project.customer_t C
INNER JOIN
    gl_project.order_t O
ON
    C.customer_id = O.customer_id
GROUP BY
    C.credit_card_type
ORDER BY
    avg_discount DESC;
```

Output:

credit_card_type	avg_discount
laser	0.643846
mastercard	0.629500
maestro	0.624219
visa-electron	0.623469
china-unionpay	0.622174
instapayment	0.620625
americanexpress	0.616327
diners-club-us-ca	0.614615
diners-club-carte-blanche	0.614490
switch	0.610233
bankcard	0.609545
jcb	0.607382
visa	0.600833
diners-club-enroute	0.599792
solo	0.585000
diners-club-international	0.584000

Graph:



Observations and Insights:

- **Credit Card Discounts**:
  - Laser, Mastercard, Maestro, and Visa-Electron offer the highest discounts.
- **Discount Range**:
  - Most average discounts lie between .60 and .64.
- **Quarterly Trends**:
  - Q3 had the highest average discount (.69), and Q1 the lowest (.54).
  - Q4's discount was .63.
- **Impact of Discounts**:
  - Higher discounts (post-Q1) have not led to higher sales, suggesting a need to reassess discount strategies.

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

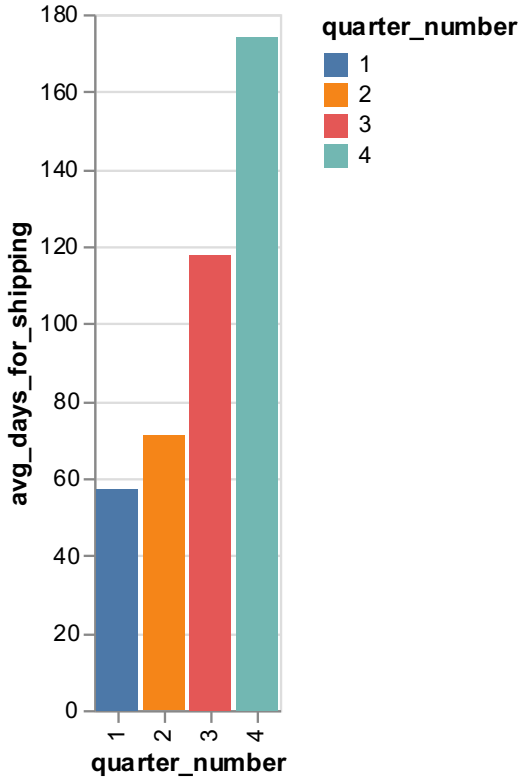
Solution Query:

```
SELECT
    quarter_number,
    AVG(DATEDIFF(ship_date, order_date)) AS avg_days_for_shipping
FROM
    gl_project.order_t
GROUP BY
    quarter_number
ORDER BY
    quarter_number ASC;
```

Output:

quarter_numb...	avg_days_for_shippi...
1	57.1677
2	71.1107
3	117.7555
4	174.0955

Graph:



Observations:

- There is trend of increasing delivery time from Q1 to Q4
- This might have made customer frustrated which shows influx of so many negative reviews on later part of year.
- Deeper analysis required to identify factor which influence the customer satisfaction.

Business Metrics Overview

Total Revenue	Total Orders	Total Customers	Average Rating
124.71 Million	1000	994	3.14
Last Quarter Revenue	**Last quarter Orders**	**Average Days to Ship**	**% Good Feedback**
23.35 Million	199	97.96	21.5

Business Recommendations

- **Vehicle Makers:**
  - As Chevrolet and Ford are most demanded vehicle try to cash in the demand for Chevrolet and Ford by maintaining a sufficient inventory and offering competitive price.
  - Develop targeted approaches to elevate customer engagement and satisfaction for Toyota, Pontiac, and Dodge to boost their market share and strengthen their positions
- **Customer Ratings:**
  - Customer ratings are quite less in Q4 and there is a decreasing trend. Analyze customer rating and find the reason for declining rating.
  - We had seen even Delivery time is increased from Q1 to Q4. That might have frustrated the user as well which results lower feedback

- **No Of Orders and Revenue:**
  - There is a decrease in the order from Q1 to Q4. And this might be due to experience people have from new wheels.
  - Net Revenue is steeply down in Q4 which might be due to decrease in orders/sales. But we need to analyze more on discount to get clear idea if there is any effect of discount as well on this.

Recommendation:

- Identify patterns in delivery delays, especially during peak periods like holidays or sales events. Preemptively scale logistics and workforce to handle surges in order volume. Optimize Supply Chain and Logistics and partner with reliable and efficient shipping providers to improve delivery speed and reduce delays.
- Try to cash in the demand for Chevrolet and Ford by maintaining a sufficient inventory and offering competitive price.
- Reengage previous customers through personalized email campaigns or targeted ads.
- Listen to customer feedback and working on those will improve satisfaction.
- Setup Post sales service to improve the experience of customer.