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.

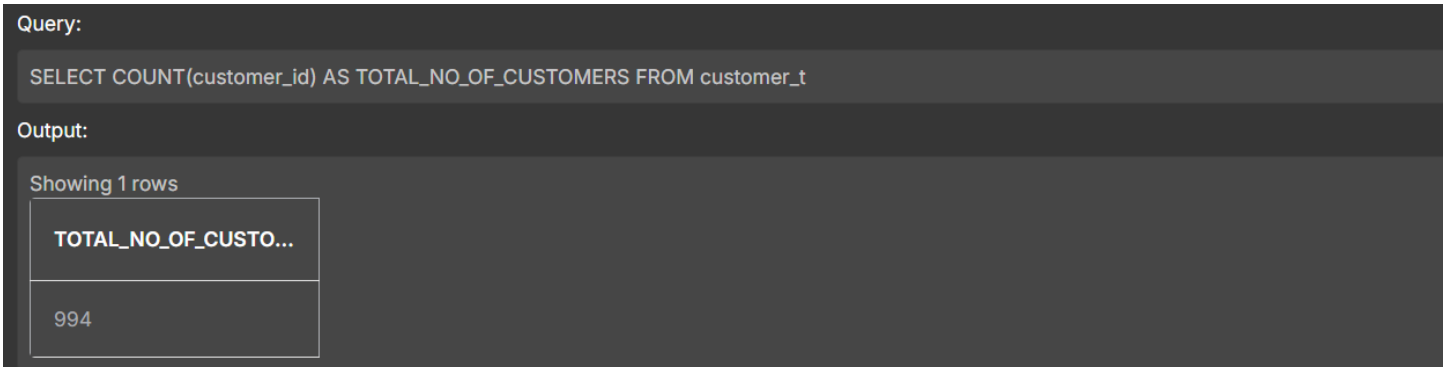
Business Questions

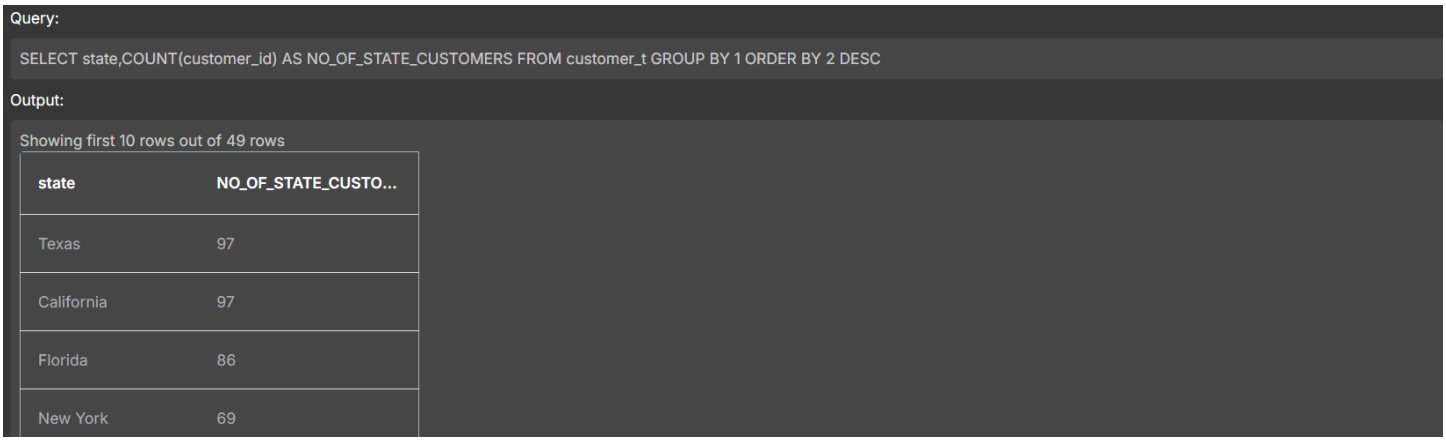
Question 1: Find the total number of customers who have placed orders. What is the distribution of the customers across states?

Solution Query:

```
SELECT
        COUNT(customer_id) AS TOTAL_NO_OF_CUSTOMER
FROM customer_t;
SELECT
        state,
        COUNT(customer_id) AS NO_OF_STATE_CUSTOMERS
FROM customer_t
GROUP BY 1
ORDER BY 2 DESC;
```

Output:





Observations and Insights:

- All above observation, A total of 994 customers has placed their orders.

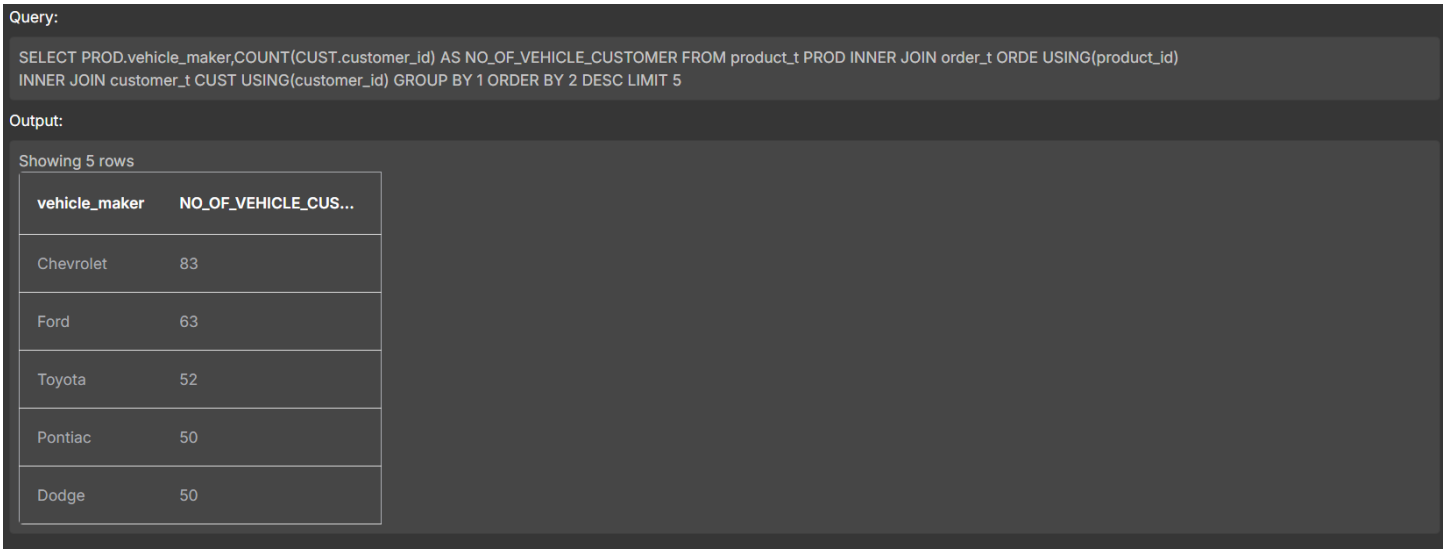
- The highest number of customers are from Texas and California followed by Florida, New Yark and the District of Columbia.

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

Solution Query:

```
SELECT
    PROD.vehicle_maker,
    COUNT(CUST.customer_id) AS NO_OF_VEHICLE_CUSTOMER
FROM product_t PROD
INNER JOIN order_t ORDE USING(product_id)
INNER JOIN customer_t CUST USING(customer_id)
GROUP BY 1
ORDER BY 2 DESC
LIMIT 5;
```

Output:



Observations and Insights:

- The most five favored vehicle manufacturers among customers are Chevrolet, Ford, Toyota, Pontiac and Dodge.
- Chevrolet is the most favored vehicle manufacturer among customers, followed by Ford and Toyota coming next.

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

Solution Query:

```
SELECT
        CUST.state,
        PROD.vehicle_maker
FROM product_t AS PROD
INNER JOIN order_t AS ORDE USING(product_id)
INNER JOIN customer_t AS CUST USING(customer_id)
GROUP BY 1
ORDER BY COUNT(CUST.customer_id) DESC;
```

Output:

Query:

```
SELECT CUST.state,PROD.vehicle_maker FROM product_t AS PROD INNER JOIN order_t AS ORDE USING(product_id)
INNER JOIN customer_t AS CUST USING(customer_id) GROUP BY 1 ORDER BY COUNT(CUST.customer_id) DESC
```

Output:

Showing first 10 rows out of 49 rows

state	vehicle_maker
Texas	Chevrolet
California	Toyota
Florida	GMC
New York	Plymouth
District of Columbia	Dodge

Observations and Insights:

- The above observation said that the vehicle preferences of customers have a lot of variability by state.
- Chevrolet is the most popular vehicle maker in the state of Texas, followed by Toyota, GMC, Plymouth and Dodge in the states of California, Florida, New York, and the District of Columbia, respectively.

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
        ROUND(AVG(CUST_FEEDBACK),2) AS AVG_FEED_BACK
FROM (
        SELECT
                CASE
                WHEN customer_feedback='Very Good' then 5
                WHEN customer_feedback='Good' then 4
                WHEN customer_feedback='Okey' then 3
                WHEN customer_feedback='Bed' then 2
                 WHEN customer_feedback='Very Bed' then 1
                END AS CUST_FEEDBACK
        FROM order_t
        );
SELECT
        quarter_number,
        ROUND(AVG(CUST_FEEDBACK),2) AS AVG_FEED_BACK
FROM (
        SELECT
                quarter_number,
                CASE
                WHEN customer_feedback='Very Good' then 5
```

```
        WHEN customer_feedback='Good' then 4
        WHEN customer_feedback='Okey' then 3
        WHEN customer_feedback='Bed' then 2
        WHEN customer_feedback='Very Bed' then 1
        END AS CUST_FEEDBACK
    FROM order_t
    ) AS FEEDBACK_BUCK
GROUP BY 1
ORDER BY 1;
```

Output:

Query:

```
SELECT ROUND(AVG(CUST_FEEDBACK),2) AS AVG_FEED_BACK FROM (SELECT CASE  WHEN customer_feedback='Very Good' then 5 WHEN customer_feedback='Good' then 4
WHEN customer_feedback='Okey' then 3 WHEN customer_feedback='Bed' then 2 WHEN customer_feedback='Very Bed' then 1
END AS  CUST_FEEDBACK FROM order_t)
```

Output:

Showing 1 rows

AVG_FEED_BACK
4.51

Query:

```
SELECT quarter_number, ROUND(AVG(CUST_FEEDBACK),2) AS AVG_FEED_BACK FROM (SELECT quarter_number,CASE  WHEN customer_feedback='Very Good' then 5
WHEN customer_feedback='Good' then 4 WHEN customer_feedback='Okey' then 3 WHEN customer_feedback='Bed' then 2 WHEN customer_feedback='Very Bed' then 1
END AS CUST_FEEDBACK FROM order_t) AS FEEDBACK_BUCK GROUP BY 1 ORDER BY 1
```

Output:

Showing 4 rows

quarter_number	AVG_FEED_BACK
1	4.51
2	4.56
3	4.44
4	4.5

Observations and Insights:

- The overall average feedback is 4.51, which is good average feedback.
- The average quarterly feedback is between 4.4 to 4.6. There is not a big quarterly rating difference in customers' feedback.

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 Good' then 1 ELSE 0 END)*
        100/COUNT(customer_feedback)),2) AS PERCENT_VERY_GOOD,
         ROUND((SUM(CASE WHEN customer_feedback='Good' then 1 ELSE 0 END)*
        100/COUNT(customer_feedback)),2) AS PERCENT_GOOD,
        ROUND((SUM(CASE WHEN customer_feedback='Okey' then 1 ELSE 0 END)*
        100/COUNT(customer_feedback)),2) AS PERCENT_OKEY,
        ROUND((SUM(CASE WHEN customer_feedback='bed' then 1 ELSE 0 END)*
        100/COUNT(customer_feedback)),2) AS PERCENT_BED,
        ROUND((SUM(CASE WHEN customer_feedback='Very bed' then 1 ELSE 0 END)*
        100/COUNT(customer_feedback)),2) AS PERCENT_VERY_BED
FROM order_t
GROUP BY 1
ORDER BY 1;
```

Output:

Query:

```
SELECT quarter_number,ROUND((SUM(CASE WHEN customer_feedback='Very Good' then 1 ELSE 0 END)*100/COUNT(customer_feedback)),2) AS PERCENT_VERY_GOOD,
ROUND((SUM(CASE WHEN customer_feedback='Good' then 1 ELSE 0 END)*100/COUNT(customer_feedback)),2) AS PERCENT_GOOD,
ROUND((SUM(CASE WHEN customer_feedback='Okey' then 1 ELSE 0 END)*100/COUNT(customer_feedback)),2) AS PERCENT_OKEY,
ROUND((SUM(CASE WHEN customer_feedback='bed' then 1 ELSE 0 END)*100/COUNT(customer_feedback)),2) AS PERCENT_BED,
ROUND((SUM(CASE WHEN customer_feedback='Very bed' then 1 ELSE 0 END)*100/COUNT(customer_feedback)),2) AS PERCENT_VERY_BED
FROM order_t GROUP BY 1 ORDER BY 1
```

Output:

Showing 4 rows

quarter_number	PERCENT_VERY_GOOD	PERCENT_GOOD	PERCENT_OKEY	PERCENT_BED	PERCENT_VERY_BED
1	30	28	0	0	0
2	28	22	0	0	0
3	16	20	0	0	0
4	10	10	0	0	0

Observations and Insights:

- As shown above, customer satisfaction has gone down over time therefore customers are getting dissatisfied over time.
- in Quarter 1, the ratings are relatively good with over half of the customers expressing satisfaction.
- However, the ratings decline significantly in subsequent quarters.
  In Quarter 4, the ratings show that over half of the customers are highly dissatisfied.

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

Solution Query:

```
SELECT
    quarter_number,
    COUNT(order_id) AS QUARTER_ORDERS_COUNT
FROM order_t
GROUP BY 1
ORDER BY 2;
```

Output:

Query:

```
SELECT quarter_number,COUNT(order_id) AS QUARTER_ORDERS_COUNT FROM order_t GROUP BY 1 ORDER BY 2
```

Output:

Showing 4 rows

quarter_number	QUARTER_ORDERS_C...
4	199
3	229
2	262
1	310

Observations and Insights:

- Customers placed the highest number of orders in Quarter 1, but Customer orders decreased by a quarter on quarter subsequently.
- With both, the number of orders and customer satisfaction is going down. The company needs improvement.

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

Solution Query:

```
SELECT
        ROUND(SUM(REV),2) AS TOTAL_REVENUE
FROM (
        SELECT
                quarter_number,
                SUM(quantity*(vehicle_price-(vehicle_price*(discount/100)))) AS REV
        FROM order_t) AS REVNUE;


SELECT
        quarter_number,
        REV,
        LAG(REV) OVER(ORDER BY quarter_number) AS QUARTER_PREVIOUS_REVENUE,
        (REV-LAG(REV) OVER(ORDER BY quarter_number))/
                LAG(REV) OVER(ORDER BY quarter_number) AS  QUARTER_PERCENT_CHANGE
FROM (
        SELECT
                 quarter_number,
                SUM(quantity*(vehicle_price-(vehicle_price*(discount/100)))) AS REV
        FROM order_t
        GROUP BY quarter_number
         ) AS QUARTER_REVENUE
ORDER BY 1;
```

Output:

Query:

```
SELECT ROUND(SUM(REV),2) AS TOTAL_REVENUE FROM (SELECT quarter_number,SUM(quantity*(vehicle_price-(vehicle_price*(discount/100)))) AS REV FROM order_t) AS REVNUE
```

Output:

Showing 1 rows

TOTAL_REVENUE
124714086.32

Query:

```
SELECT quarter_number,REV,LAG(REV) OVER(ORDER BY quarter_number) AS QUARTER_PREVIOUS_REVENUE,
  (REV-LAG(REV) OVER(ORDER BY quarter_number))/LAG(REV) OVER(ORDER BY quarter_number) AS  QUARTER_PERCENT_CHANGE_REV
  FROM (SELECT quarter_number,SUM(quantity*(vehicle_price-(vehicle_price*(discount/100)))) AS REV FROM order_t  GROUP BY quarter_number) AS QUARTER_REVENUE ORDER BY 1
```

Output:

Showing 4 rows

quarter_number	REV	QUARTER_PREVIOUS_...	QUARTER_PERCENT_...
1	39421580.15929598		
2	32715830.33996199	39421580.15929598	-0.17010352685603125
3	29229896.19364898	32715830.33996199	-0.10655190805458437
4	23346779.63060599	29229896.19364898	-0.20127052535757087

Observations and Insights:

- The total revenue generated over the year is ~ $124.7 M.
- There is a significant revenue decline of approximately 59% from Quarter 1 to Quarter 4, which is quite unfavorable.
- The above observations correlate with previous ones as with the number of orders going down, revenue generated has decreased

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

Solution Query:

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

Output:

Query:

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

Output:

Showing 4 rows

quarter_number	REVENUE	COUNT_QUARTER_OR...
1	39421580.15929598	310
2	32715830.33996199	262
3	29229896.19364898	229
4	23346779.63060599	199

Observations and Insights:

- The 1st quarter has higher net revenue followed by 2nd quarter, 3rd quarter and 4th quarter.
- The Number of orders has decreased, with revenue per quarter declining.

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

Solution Query:

```
SELECT
    CUST.credit_card_type,
    ROUND(AVG(ORDE.discount),2) AVG_CREDIT_CARD_DISCOUNT
FROM customer_t AS CUST
INNER JOIN order_t AS ORDE USING(customer_id)
GROUP BY 1
ORDER BY 2 DESC;
```

Output:

Query:

```
SELECT CUST.credit_card_type, ROUND(AVG(ORDE.discount),2) AVG_CREDIT_CARD_DISCOUNT
  FROM customer_t AS CUST INNER JOIN order_t AS ORDE USING(customer_id) GROUP BY 1
  ORDER BY 2 DESC
```

Output:

Showing first 10 rows out of 16 rows

credit_card_type	AVG_CREDIT_CARD_DI...
laser	0.64
mastercard	0.63
visa-electron	0.62
maestro	0.62
instapayment	0.62

Observations and Insights:

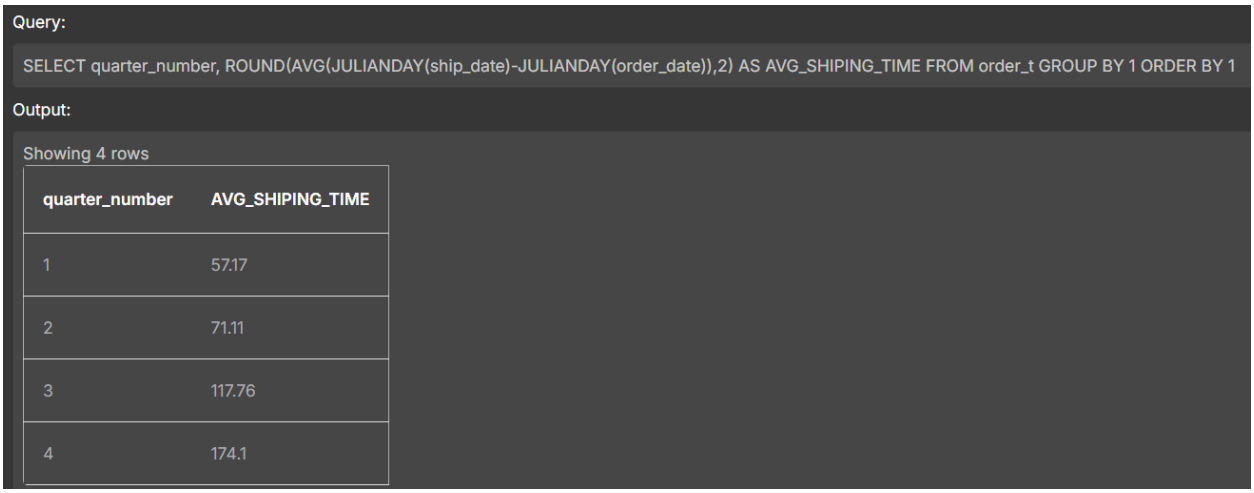
- The discount offered varies between ~58% to ~64%
- Laser has offered 64.38% of the highest average discount to customers. And the lowest average discount offered by Diners-club-international

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 AVG_SHIPING_TIME
FROM order_t
GROUP BY 1
ORDER BY 1;
```

Output:



Observations and Insights:

- As shown above, Overall, the average shipping time has increased from the 1st quarter to the 4th quarter.

- The 1st quarter average shipping time was 57 days, and the 4th quarter average shipping time was 174 days. The difference between the 1st quarter and the 4th quarter is 117 days. Which is undesirable for the customers.

Business Metrics Overview

Total Revenue	Total Orders	Total Customers	Average Rating
124.7M	100	994	4.51
Last Quarter Revenue	**Last quarter Orders**	**Average Days to Ship**	**% Good Feedback**
23.3M	199	98	20%

Business Recommendations

- The Company must consider shipping time because the shipping time has increased nearly 3 times in a single year. Hiring more shipping and delivery executives might aid in resolving this issue.
- The Company lost nearly 59% of its net revenue in one year due to decreasing orders. The Company must take the necessary steps to identify the underlying reason for this decline.
- It is recommended that a campaign gathers customer preference for vehicles for the respective regions. The Company can start the test campaign in Texas, California, Florida, and New York.
- It must recommend working on customer services and satisfaction. The overall average feedback is good. The percentage distribution of feedback is decreasing from customers.