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:

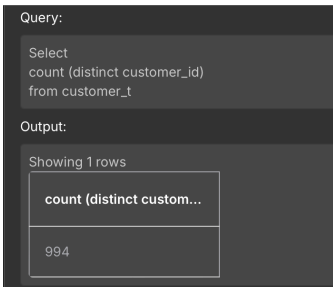
Total number of customers who placed orders:

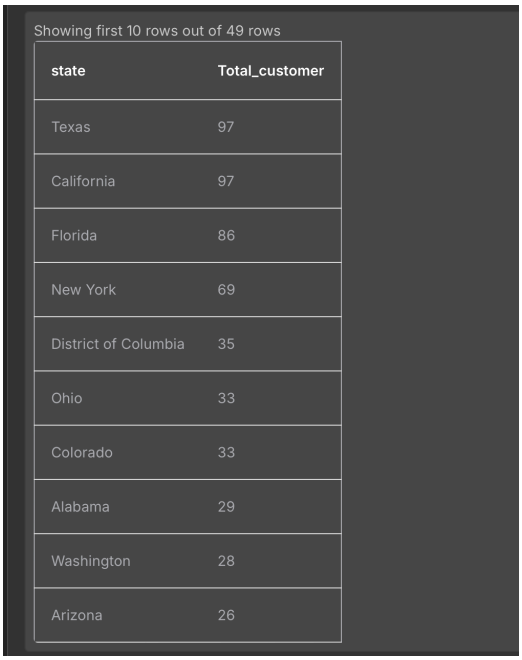
```
Select
count (distinct customer_id)
from customer_t;
```

Distribution of the customers across states:

```
Select
c.state,
count ( distinct o.customer_id) as Total_customer
from customer_t c
join order_t o
on o.customer_id=c.customer_id
group by 1
order by 2 desc;
```

Output:





state	Total_customer
Texas	97
California	97
Florida	86
New York	69
District of Columbia	35
Ohio	33
Colorado	33
Alabama	29
Washington	28
Arizona	26

Observations and Insights:

- **The total count of customers was 994**
- **The states with the most customers are Texas 97, California 97, Florida 86.**
- **Find ways to continue to target higher selling states. What discounts can be incorporated?**
- **What type of marketing/ discounts can happen in lower selling states? Those states may have as many customers for various reasons, but what are somethings that can happen so there are more customers in each state?**
- **How are the customer totals in comparison to the total population?**

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

Solution Query:

```
Select
p.vehicle_maker,
c.customer_id,
count (o.customer_id) as total_vehicle_maker

from customer_t c
      join order_t o
            on o.customer_id=c.customer_id
      join product_t p
            on p.product_id=o.product_id
group by 1
order by 3 desc
limit 5;
```

Output:

Showing 5 rows

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

Observations and Insights:

- **The top vehicles are Chevrolet, Ford, Toyota, Pontiac, Dodge**
- **Why are some specifics that make the vehicles the highest amount sold?**
- **Where are these vehicles mostly sold?**

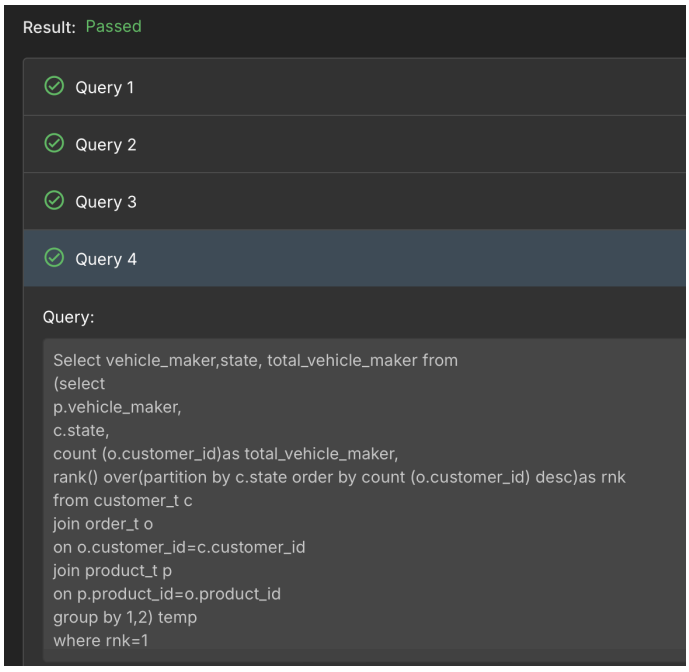
Question 3: Which is the most preferred vehicle maker in each state?

Solution Query:

```
Select vehicle_maker,state, total_vehicle_maker

from
(select
      p.vehicle_maker,
      c.state,
count (o.customer_id)as total_vehicle_maker,
rank() over(partition by c.state order by count (o.customer_id) desc)as rnk
from customer_t c
      join order_t o
            on o.customer_id=c.customer_id
      join product_t p
            on p.product_id=o.product_id
group by 1,2) temp
where rnk=1;
```

Output:



vehicle_maker	state	total_vehicle_maker
Dodge	Alabama	5
Chevrolet	Alaska	2
Cadillac	Arizona	3
Pontiac	Arizona	3
Chevrolet	Arkansas	1
GMC	Arkansas	1
Mitsubishi	Arkansas	1
Pontiac	Arkansas	1
Suzuki	Arkansas	1
Volkswagen	Arkansas	1

Observations and Insights:

- **Dodge was the most preferred vehicle in Alabama**
- **Chevrolet was the most preferred vehicle in Alaska, Arizona's most preferred vehicle was Pontiac**
- **Arkansas had a variety of vehicles with only 1 sold**

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:

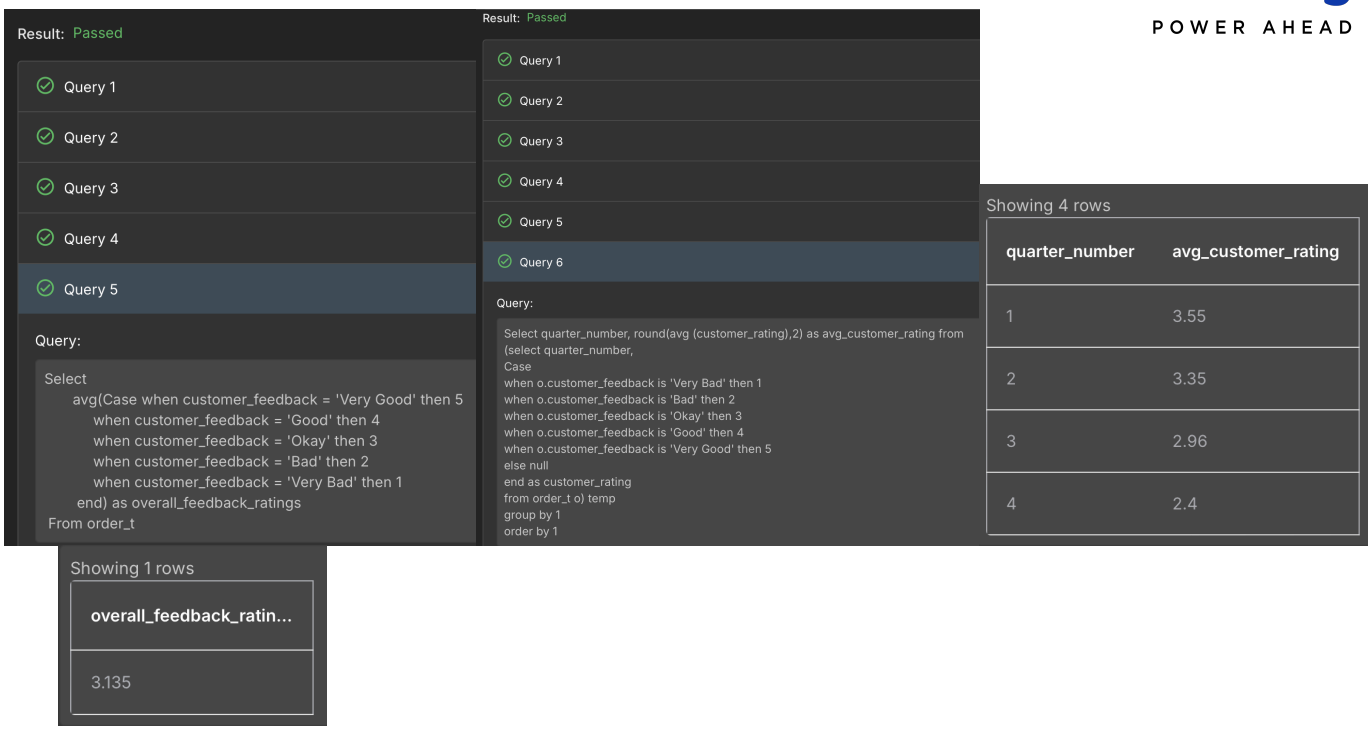
Overall Average Rating

```
Select
       avg(Case
            when customer_feedback = 'Very Good' then 5
            when customer_feedback = 'Good' then 4
            when customer_feedback = 'Okay' then 3
            when customer_feedback = 'Bad' then 2
            when customer_feedback = 'Very Bad' then 1
        end) as overall_feedback_ratings
 From order_t;
```

Average rating for each quarter

```
Select quarter_number, round(avg (customer_rating),2) as avg_customer_rating
from
      (select quarter_number,
      Case
            when o.customer_feedback is 'Very Bad' then 1
            when o.customer_feedback is 'Bad' then 2
            when o.customer_feedback is 'Okay' then 3
            when o.customer_feedback is 'Good' then 4
            when o.customer_feedback is 'Very Good' then 5
            else null
            end as customer_rating
            from order_t o) temp
group by 1
order by 1
```

Output:



Observations and Insights:

- **The average customer rating is trending in a negative direction**
- **Quarter 1 had the highest customer rating**
- **Quarter 4 had the lowest customer rating, the customer ratings began to trend so poorly that the overall customer ratings aren't even as high as Quarter 1 or Quarter**

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

Solution Query:

```
Select quarter_number,
round(very_bad_count*100.0/total_feedback,2) as perc_very_bad,
round(bad_count*100.0/total_feedback,2) as perc_bad_count,
round(okay_count*100.0/total_feedback,2) as perc_okay_count,
round(good_count*100.0/total_feedback,2) as perc_good_count,
round(very_good_count*100.0/total_feedback,2) as perc_very_good_count
from
(SELECT
    quarter_number,
    SUM(CASE WHEN customer_feedback = 'Very Bad' THEN 1 ELSE 0 END) AS
very_bad_count,
    SUM(CASE WHEN customer_feedback = 'Bad' THEN 1 ELSE 0 END) AS bad_count,
    SUM(CASE WHEN customer_feedback = 'Okay' THEN 1 ELSE 0 END) AS okay_count,
    SUM(CASE WHEN customer_feedback = 'Good' THEN 1 ELSE 0 END) AS good_count,
    SUM(CASE WHEN customer_feedback = 'Very Good' THEN 1 ELSE 0 END) AS
very_good_count,
    count (customer_feedback) as total_feedback
FROM order_t
GROUP BY quarter_number)
;
```

Output:

Result: Passed

Query 1

Query:

```
Select quarter_number,
round(very_bad_count*100.0/total_feedback,2) as perc_very_bad,
round(bad_count*100.0/total_feedback,2) as perc_bad_count,
round(okay_count*100.0/total_feedback,2) as perc_okay_count,
round(good_count*100.0/total_feedback,2) as perc_good_count,
round(very_good_count*100.0/total_feedback,2) as perc_very_good_count
from
(SELECT
   quarter_number,
   SUM(CASE WHEN customer_feedback = 'Very Bad' THEN 1 ELSE 0 END) AS very_bad_count,
   SUM(CASE WHEN customer_feedback = 'Bad' THEN 1 ELSE 0 END) AS bad_count,
   SUM(CASE WHEN customer_feedback = 'Okay' THEN 1 ELSE 0 END) AS okay_count,
   SUM(CASE WHEN customer_feedback = 'Good' THEN 1 ELSE 0 END) AS good_count,
   SUM(CASE WHEN customer_feedback = 'Very Good' THEN 1 ELSE 0 END) AS very_good_count,
   count (customer_feedback) as total_feedback
FROM order_t
GROUP BY quarter_number)
```

Output:

Showing 4 rows

quarter_number	perc_very_bad	perc_bad_count	perc_okay_count	perc_good_count	perc_very_good_count
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:

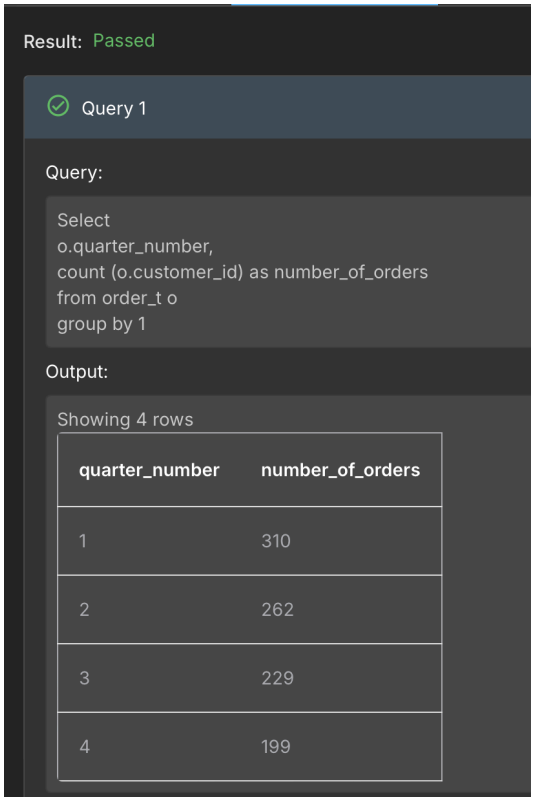
- **As you go through each quarter the percentage of " Very Good" is cut by a third**
- **Very Bad percentage feed back nearly triples**
- **Customers are clearly getting dissatisfied as time goes on.**

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

Solution Query:

```
Select
      o.quarter_number,
      count (o.customer_id) as number_of_orders
from order_t o
group by 1;
```

Output:



Observations and Insights:

- **Quarter 1 had the highest amount of sales**
- **As time goes on the quarterly sales decrease**
- **Biggest drop off was between Q1 and Q2 for orders**

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

Solution Query:

```
Select *,
      round((revenue - lag(revenue,1) over(order by quarter_number) )/
      lag(revenue,1) over(order by quarter_number),2) as perc_qoq
from
      (Select quarter_number,
      sum(quantity * (vehicle_price - ((discount/100)*vehicle_price))) as
      revenue
from order_t
group by 1);
```

Output:

Result: Passed

Query 1

Query:

```
Select *,
    round((revenue - lag(revenue,1) over(order by quarter_number) )/lag(revenue,1) over(order by quarter_number),2) as perc_qoq
from
(Select quarter_number,
    sum(quantity * (vehicle_price - ((discount/100)*vehicle_price))) as revenue

from order_t
group by 1)
```

Output:

Showing 4 rows

quarter_number	revenue	perc_qoq
1	39421580.15929598	
2	32715830.33996199	-0.17
3	29229896.19364898	-0.11
4	23346779.63060599	-0.2

Observations and Insights:

- **Even though the number order drop off was greatest between Q1 and Q2, the biggest difference in revenue was between Q3 and Q4.**
- **Q1 Had greatest revenue**
- **Overtime the revenue decreased, a clear reflection of customer dissatisfaction.**

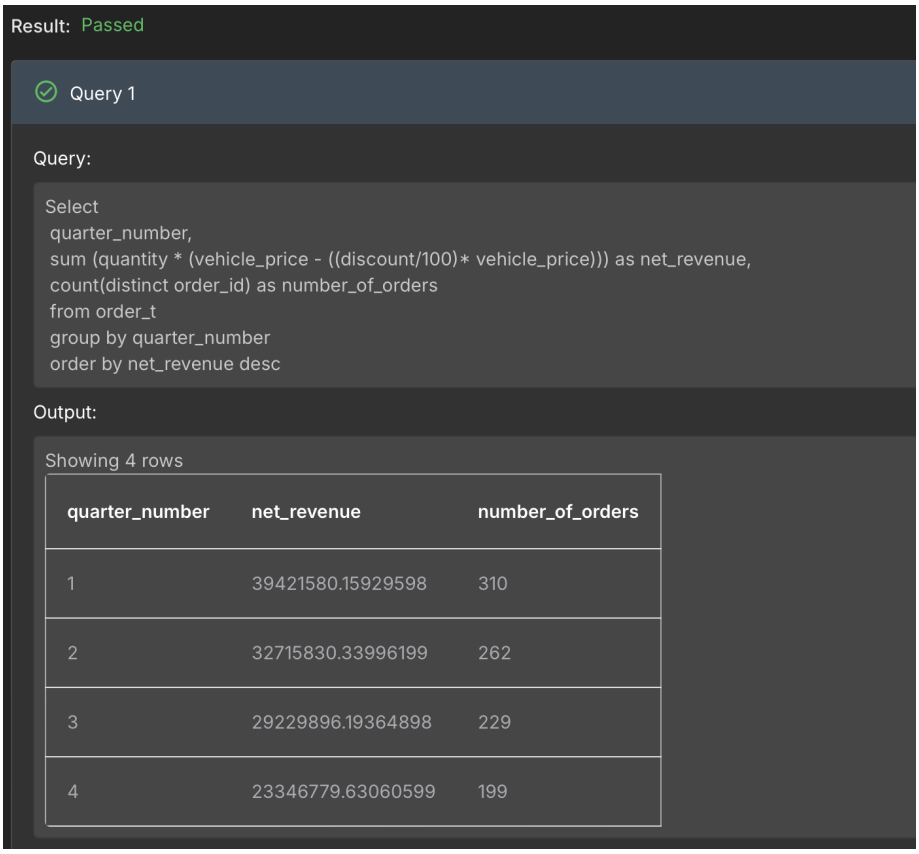
Question 8: What is the trend of net revenue and orders by quarters?

Solution Query:

```
Select
      quarter_number,
      sum (quantity * (vehicle_price - ((discount/100)* vehicle_price))) as
            net_revenue,
      count(distinct order_id) as number_of_orders

from order_t
group by quarter_number
order by net_revenue desc;
```

Output:



Observations and Insights:

- **Q1 had greatest amount of orders and highest revenue**
- **Throughout the quarters you can see a clear negative trend in both orders and net revenue**
- **Did discounts help the sales in any capacity throughout the quarters?**

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

Solution Query:

```
select
      credit_card_type,
      round(avg(discount),3) as avg_discount
      from order_t o
      join customer_t c
            on c.customer_id=o.customer_id
      group by 1
      order by 2 desc;
```

Output:

Observations and Insights:

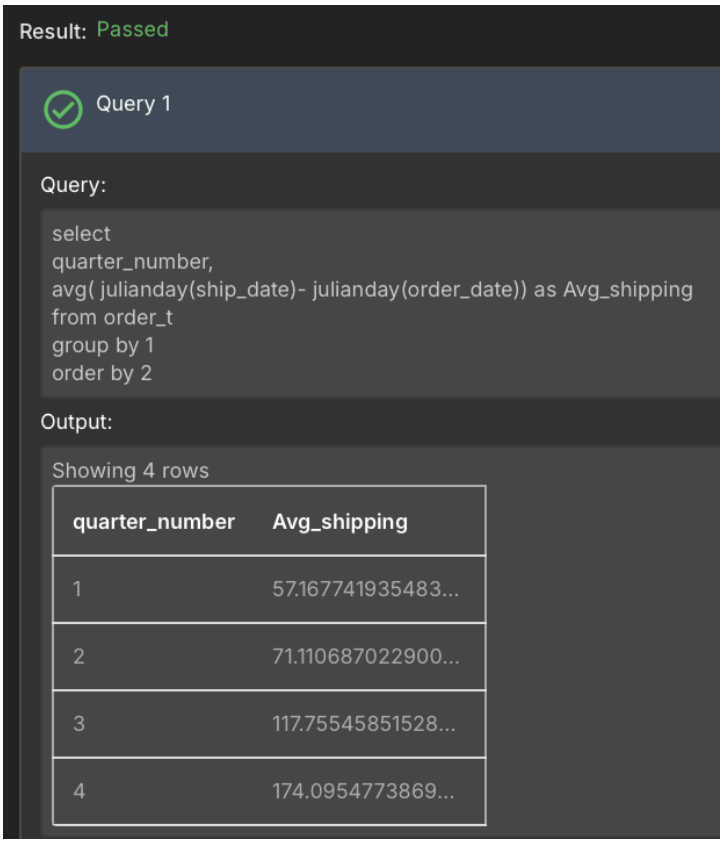
- **Laser has the best avg discount**
- **Most discounts for all Credit cards are around .6**
- **Company should may look at business opportunities from different credit unions**

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

Solution Query:

```
select
quarter_number,
avg( julianday(ship_date)- julianday(order_date)) as Avg_shipping
from order_t
group by 1
order by 2;
```

Output:



Business Metrics Overview

Total Revenue	Total Orders	Total Customers	Average Rating
$124714086.32	1000	994	3.135
Last Quarter Revenue	**Last quarter Orders**	**Average Days to Ship**	**% Good Feedback**
$23346779.63	199	97.964	21.5% - 'Good' percentage 44.1% - 'Good' + 'Very Good'

Business Recommendations

- **Find the root cause of shipping delay. This is likely the issue as to why sales and revenue dropped throughout the quarters.**
- **Discounts and promotions can be helpful, especially when there are shipping delays that can not be avoided. What types of promotions & discounts could company offer to make buying a vehicle from new wheels more enticing.**
- **What type of advertisements have been used throughout each Quarter ? It would be worth looking into spending money in advertisements to generate more sales.**