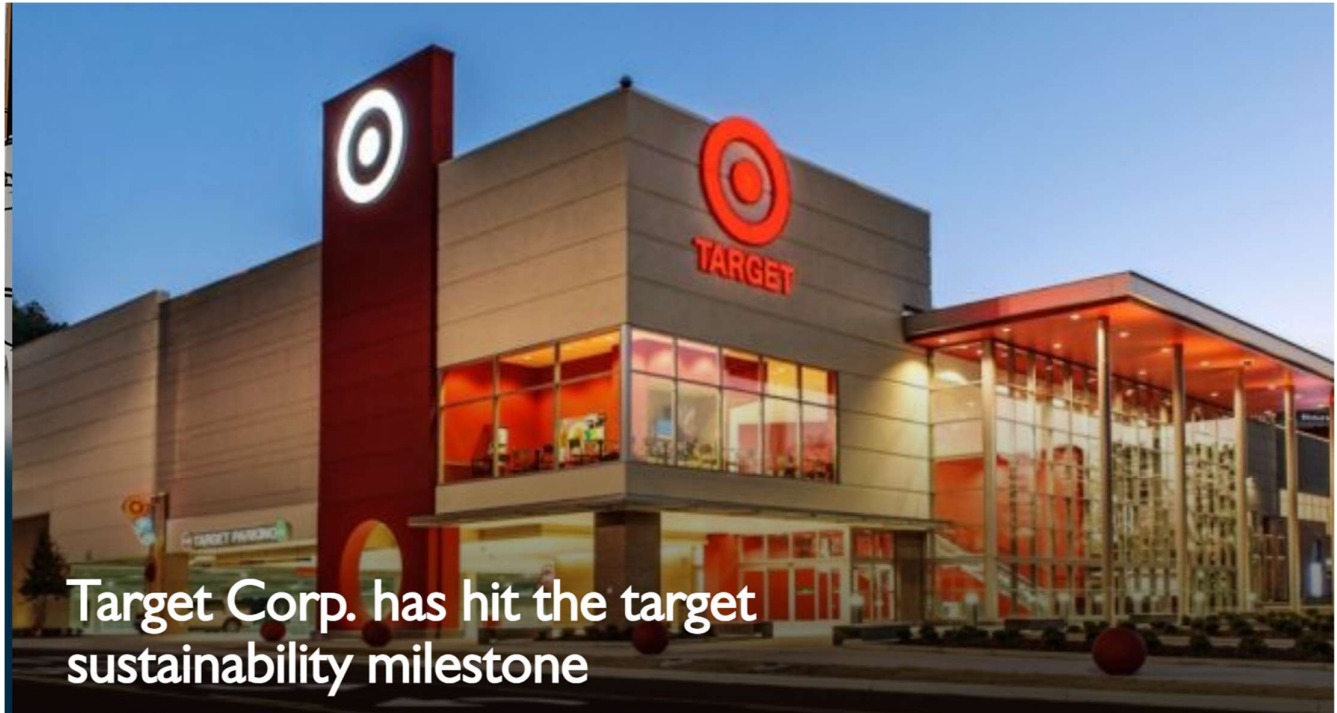


Target-Company-Business-CaseStudy-Analysis



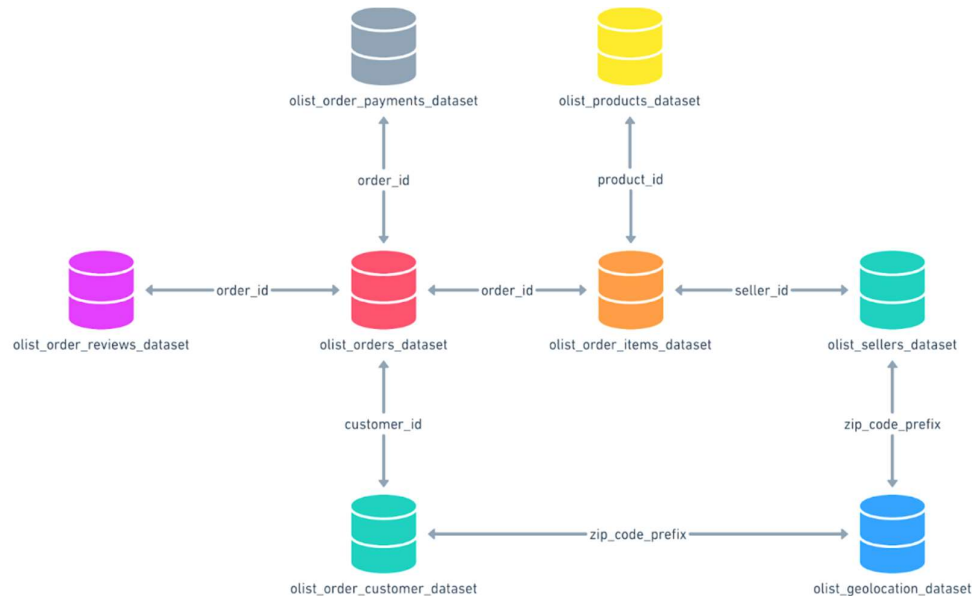
Target Corporation, headquartered in Minneapolis, is a major American retail company and the seventh-largest retailer in the US. Established in 1962 as Dayton's discount division, it expanded nationwide in the 1980s and introduced various store formats under the Target brand in the 1990s. The company is known for its cheap-chic image. It became Target Corporation in 2000 and divested department store chains in 2004. Despite challenges like security breaches and the failure of Target Canada, it found success with urban market expansion. As of 2023, Target operates 1,948 stores in the US and ranks 32 on the Fortune 500 list.

Target Corporation's history began in 1902 as Goodfellow Dry Goods. The first Target store opened in 1962, and it became Target Corporation in 2000.

This business case examines Target's operations in Brazil, analysing 100,000 orders from 2016 to 2018. It provides insights into order processing, pricing,

payment, shipping, customer demographics, product attributes, and satisfaction levels.

About Database :



Solving Business Case Study:

1. Import the dataset and do usual exploratory analysis steps like checking the structure & characteristics of the dataset:

1.a) Data type of all columns in the "customers" table.

Intuition:

Finding the data types of the columns in the customer's table.

Syntax:

```
select column_name, Data_type from
table.INFORMATION_SCHEMA.COLUMNS
where table_name = TableName
```

Query:

```
select
column_name,
Data_type
from
`target_company_businesscase.INFORMATION_SCHEMA.C
OLUMNS`
where table_name = "customers"
```

Result:

| Row | column_name ▼ | Data_type ▼ |
|-----|--------------------------|-------------|
| 1 | customer_id | STRING |
| 2 | customer_unique_id | STRING |
| 3 | customer_zip_code_prefix | INT64 |
| 4 | customer_city | STRING |
| 5 | customer_state | STRING |

Insights:

- to find the datatype of any column use
INFORMATION_SCHEMA.COLUMNS

1.b) Get the time range between which the orders were placed

Intuition:

Finding the first order placed by customers and last order placed by customers.

Query:

```

select
-- Aggregate min and max functions
min(order_purchase_timestamp) as
First_Order_Date,
max(order_purchase_timestamp) as
Last_Order_Date
from `target_company_businesscase.orders`

```

Result:

| Row | First_Order_Date ▼ | Last_Order_Date ▼ |
|-----|-------------------------|-------------------------|
| 1 | 2016-09-04 21:15:19 UTC | 2018-10-17 17:30:18 UTC |

Insights:

- The first order placed on 2016-09-04 around 9:15 PM (UTC)
- The last order placed on 2018-10-17 around 5:30 PM(UTC)

1.c) Count the Cities & States of customers who ordered during the given period

Intuition :

Finding the Total number of distinct cities and states in Brazil.

Query :

```

select
-- Using Aggregate count function with
distinct
count(distinct customer_city) as
Total_Number_Of_cities,

```

```

count(distinct customer_state) as
Total_Number_Of_states
-- fetching data from customers table
from `target_company_businesscase.customers`

```

Result :

| Row | Total_Number_Of_cities | Total_Number_Of_states |
|-----|------------------------|------------------------|
| 1 | 4119 | 27 |

Insights :

- There are 4119 total number of cities in Brazil
- There are 27 states in Brazil



- There are 5 Rich States in Brazil: Sao Paulo, Distrito Federal, Rio de Janeiro, Minas Gerais, and Espírito Santo
- There are 22 middle and poor states in Brazil.

Recommendations :

- If you want to launch large supermarkets and costly products, you can launch them in the five richest states according to cities.

2. In-depth Exploration:

2a) Is there a growing trend in the no. of orders placed over the past years?

Intuition:

We need to check if there is any growing trend in a month over one year when compared to the previous year.

Query:

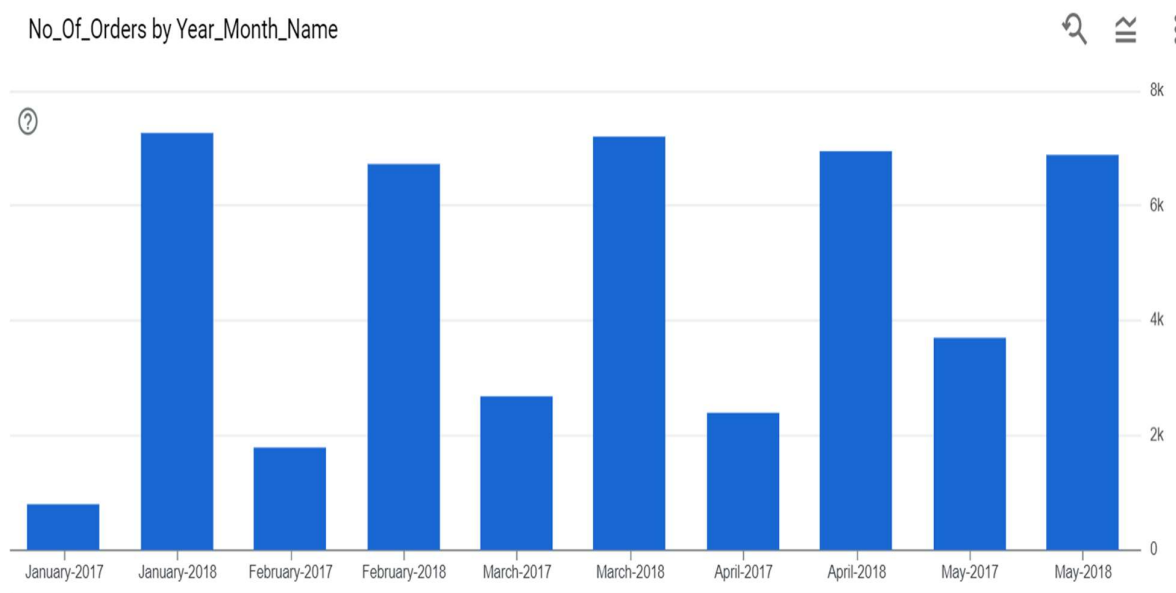
```
with CTE as (  
    select  
        extract(month from  
order_purchase_timestamp) as Month,  
        extract(year from  
order_purchase_timestamp) as Year,  
        order_id  
    from `target_company_businesscase.orders`  
)  
  
select  
Month,Year,  
count( order_id) as No_Of_Orders  
from CTE  
group by month,year
```

```
order by month
limit 10
```

Result:

| Row | Month | Year | No_Of_Orders |
|-----|-------|------|--------------|
| 1 | 9 | 2016 | 4 |
| 2 | 10 | 2016 | 324 |
| 3 | 12 | 2016 | 1 |
| 4 | 11 | 2017 | 7544 |
| 5 | 7 | 2017 | 4026 |
| 6 | 1 | 2017 | 800 |
| 7 | 8 | 2017 | 4331 |
| 8 | 6 | 2017 | 3245 |
| 9 | 2 | 2017 | 1780 |
| 10 | 10 | 2017 | 4631 |

Graph:



Insights:

- Yes, there is a growing trend in purchases in 2018 compared to 2017
- Perhaps technologies are advancing, such as online delivery.
- That's why people are ordering more products at their own time and from their own space.

Recommendations :

- Focus on Online delivery more. According to graph the year wise technology increasing.
- Always update according to situation. Current situation is online delivery system.

2b) Can we see some kind of monthly seasonality in terms of the no. of orders being placed?

Intuition :

We need to find the monthly growing trend of orders .

Query :

```
with Seasonality as (
    select
        extract(month from
order_purchase_timestamp) as Month,
        format_date('%B',
order_purchase_timestamp) as Month_Name,
        order_id
    from
`target_company_businesscase.orders`
)
```



```

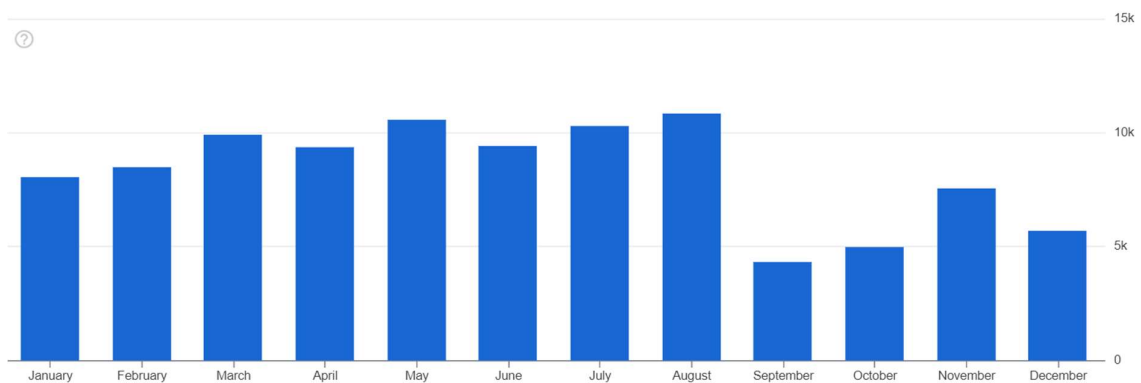
select Month_Name,Month,
       count(order_id) as No_Of_Orders
from Seasonality
group by Month_Name,Month
order by Month

```

Result :

| Row | Month_Name | Month | No_Of_Orders |
|-----|------------|-------|--------------|
| 1 | January | 1 | 8069 |
| 2 | February | 2 | 8508 |
| 3 | March | 3 | 9893 |
| 4 | April | 4 | 9343 |
| 5 | May | 5 | 10573 |
| 6 | June | 6 | 9412 |
| 7 | July | 7 | 10318 |
| 8 | August | 8 | 10843 |
| 9 | September | 9 | 4305 |
| 10 | October | 10 | 4959 |
| 11 | November | 11 | 7544 |
| 12 | December | 12 | 5674 |

Graph :



Insights :

- According to the graph, the months of March, April, May, June, and July show the highest purchases.
- This could be because Brazilian people celebrate important festivals during that time, such as Carnival in April, Festa Junina in June, Parintins Folklore Festival in July, and Oktoberfest in August .

Recommendations :

- In festival season make more discounts and offers when compare to non festival seasons

2c) During what time of the day, do the Brazilian customers mostly place their orders? (Dawn, Morning, Afternoon or Night)

■ 0-6 hrs : Dawn

■ 7-12 hrs : Mornings

■ 13-18 hrs : Afternoon

■ 19-23 hrs : Night

Intuition:

I need to find out when Brazilian customers place their orders during the (dawn, morning, afternoon, night).

Query:

```
with Div_Hour_Types as (  
select  
    case  
    when (extract(hour from  
order_purchase_timestamp) between 0 and 6)  
    then "Dawn"
```

```

when (extract(hour from
order_purchase_timestamp) between 7 and 12)
then "Mornings"
when (extract(hour from
order_purchase_timestamp) between 13 and
18) then "Afternoon"
else "Night"
end as Hour_Type,
Order_id
from `target_company_businesscase.orders`
)

```

```

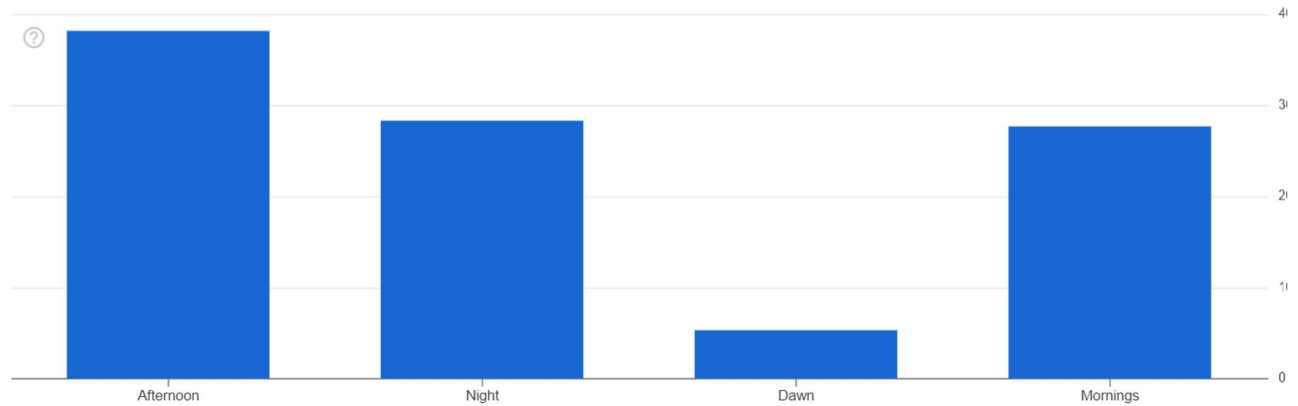
select Hour_Type, count(order_id) as
Number_Of_Orders,
dense_rank() over(order by count(order_id) desc )
as Rank
from Div_Hour_Types
group by Hour_Type

```

Result:

| Row | Hour_Type ▼ | Number_Of_Orders | Rank ▼ |
|-----|-------------|------------------|--------|
| 1 | Afternoon | 38135 | 1 |
| 2 | Night | 28331 | 2 |
| 3 | Dawn | 5242 | 4 |
| 4 | Mornings | 27733 | 3 |

Graph:



Insights:

- According to the graph, the number of orders placed in the afternoon is higher when compared to other times of the day
- This could be because more employees are free during the evening and night time.

Recommendations:

- Increase the operating hours of shops in the evening and at night.

3. Evolution of E-commerce orders in the Brazil region

3a) Get the month on month no. of orders placed in each state.

Intuition :

I need to analyse the data over month on month how many orders are placed over 27 states in Brazil.

Generate insights about which state having more orders.

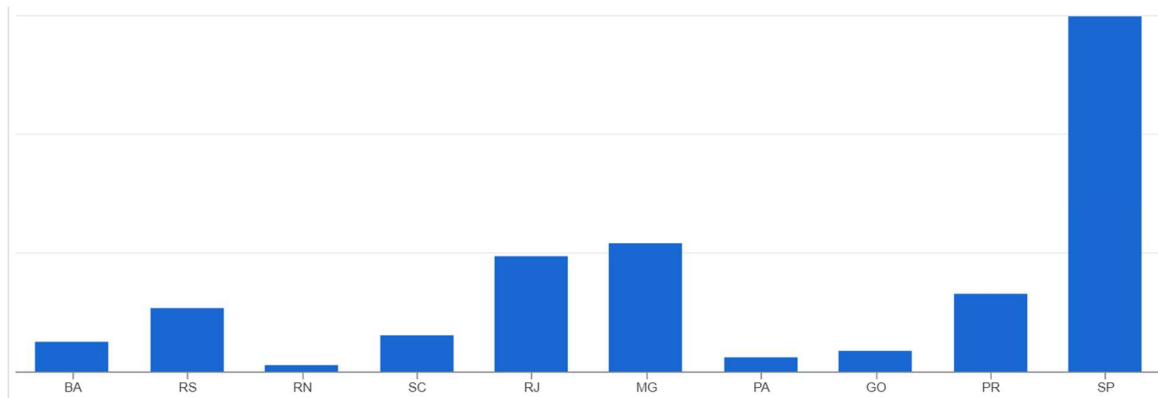
Query :

```
with Mon_On_Mon_Orders as (  
    select  
        c.customer_state,  
        extract(month from  
            o.order_purchase_timestamp) as Month,  
        extract(year from  
            o.order_purchase_timestamp) as Year,  
        format_date("%B-%Y" ,  
            o.order_purchase_timestamp) as Mon_Year,  
        o.order_id  
    from `target_company_businesscase.customers`  
    c  
    join  
        `target_company_businesscase.orders` o  
    on o.customer_id = c.customer_id  
)  
  
select  
    customer_state, Month, Year, Mon_Year,  
    count(order_id) as No_Of_Orders  
from Mon_On_Mon_Orders  
group by customer_state, Month, Year, Mon_Year  
order by Year asc  
limit 10
```

Result :

| Row | customer_state | Month | Year | Mon_Year | No_Of_Orders |
|-----|----------------|-------|------|----------------|--------------|
| 1 | CE | 10 | 2016 | October-2016 | 8 |
| 2 | RS | 10 | 2016 | October-2016 | 24 |
| 3 | SC | 10 | 2016 | October-2016 | 11 |
| 4 | RS | 9 | 2016 | September-2016 | 1 |
| 5 | RJ | 10 | 2016 | October-2016 | 56 |
| 6 | SP | 10 | 2016 | October-2016 | 113 |
| 7 | MT | 10 | 2016 | October-2016 | 3 |
| 8 | MG | 10 | 2016 | October-2016 | 40 |
| 9 | RR | 9 | 2016 | September-2016 | 1 |
| 10 | GO | 10 | 2016 | October-2016 | 9 |

Graph :



Insights :

- According to the monthly seasonality of every year, the number of orders increases or decreases.
- However, according to the graph, São Paulo registered the highest number of orders.
- Next, Minas Gerais and Rio de Janeiro registered the second and third highest number of orders in Brazil.
- The lowest number of orders were received from Rio Grande do Norte.

Recommendations:

- According to the graph, RN, PA, GO, and BA have fewer orders.
- After analysing the data, it appears that there are fewer online orders placed in these states. It is possible that there are more illiterates in these areas, and as a result, people prefer using supermarkets instead of online delivery.
- This change in behaviour contributes to increased sales for the company.

- In areas with the highest number of orders, such as SP, etc., we need to provide more costly and branded items. They will be more likely to be sold easily.

3b) How are the customers distributed across all the states?

Intuition :

We need to find how customers are spread across each state in Brazil. Based on that we can give recommendation's to increase sales of the Target.

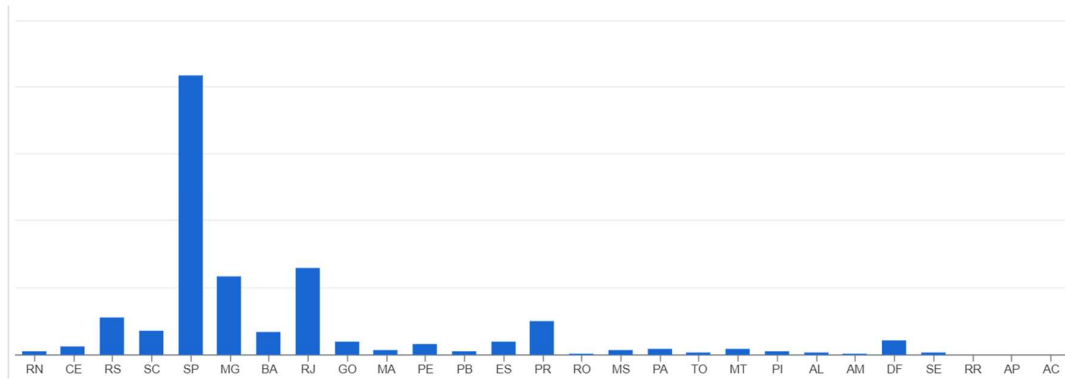
Query :

```
select
customer_state,
count(distinct customer_id) as
Number_Of_Customers
from `target_company_businesscase.customers`
group by customer_state
limit 10
```

Result :

| Row | customer_state | Number_Of_Customers |
|-----|----------------|---------------------|
| 1 | RN | 485 |
| 2 | CE | 1336 |
| 3 | RS | 5466 |
| 4 | SC | 3637 |
| 5 | SP | 41746 |
| 6 | MG | 11635 |
| 7 | BA | 3380 |
| 8 | RJ | 12852 |
| 9 | GO | 2020 |
| 10 | MA | 747 |

Graph :



Insights :

- There is the highest number of customers in São Paulo.
- The next states with similar numbers of customers are RJ, MG, RS, and PR.
- The lowest numbers of customers are in areas like AM, SE, AL, RN, and PI.

Recommendations :

- São Paulo contains the maximum number of customers, so we can target that state along with others like RJ, MG, etc.
- We can increase the number of supermarkets in these areas and offer a variety of costly and branded items.
- Providing more offers and deals, such as monthly coupons for purchases, will attract more attention from the people.
- It's essential to maintain the quality of products with high standards.

4. Impact on Economy: Analyse the money movement by e-commerce by looking at order prices, freight and others.

4a) Get the % increase in the cost of orders from year 2017 to 2018 (include months between Jan to Aug only). You can use the "payment_value" column in the payments table to get the cost of orders.

Intuition:

We need to compare the pervious year payment values to current payment values.

We need to find the how much percentage increase in payment_value in 2017 when compare to 2018.

Query:

```
with sales_betwen_2017_2018 as (  
    select  
        extract(year from  
            o.order_purchase_timestamp) as Year,  
        round(sum(p.payment_value),2) as  
            current_Payment,  
        lag(round(sum(p.payment_value),2),1)  
            over(order by sum(p.payment_value) asc)  
            as pervious_payment  
    from `target_company_businesscase.orders` o  
    join  
        `target_company_businesscase.payments` p  
    on o.order_id = p.order_id  
    where extract(year from  
        o.order_purchase_timestamp) between 2017 and  
        2018  
    and  
        extract(month from  
            o.order_purchase_timestamp) between 1 and 8  
    group by extract(year from  
        o.order_purchase_timestamp)  
)
```

```

select
    Year,
    current_payment,
    pervious_payment,
    round(((current_payment -
    pervious_payment)/pervious_payment) * 100,2)
    as percentage
from sales_betwen_2017_2018
order by Year

```

Result:

| Row | Year | current_payment | pervious_payment | percentage |
|-----|------|-----------------|------------------|-------------|
| 1 | 2017 | 3669022.12 | <i>null</i> | <i>null</i> |
| 2 | 2018 | 8694733.84 | 3669022.12 | 136.98 |

Insights:

- The current payment for 2017 is 3,669,022.12, and for 2018, it is 8,694,733.84.
- The percentage increase in 2018 compared to 2017 is 136.98%.

Recommendations:

- Investigate the factors that led to the substantial increase in payments from 2017 to 2018. Look into specific product lines, marketing campaigns, or operational improvements that may have played a role. Identifying these contributing factors can help you replicate successful strategies in the future. Investigate the factors that led to the substantial increase in payments from 2017 to 2018. Look into specific product lines, marketing campaigns, or operational improvements that may have played a role. Identifying these contributing factors can help you replicate successful strategies in the future.
- With evidence of growth in payments, consider investing more in marketing and outreach efforts. Target potential

customers through various channels, including digital marketing, social media, and traditional advertising, to expand the customer base further.

- As the company grows, maintaining operational efficiency becomes crucial. Streamline processes, optimize supply chain management, and utilize technology to improve productivity and reduce costs.
- Well-trained and motivated employees play a crucial role in providing excellent customer service and maintaining product quality. Invest in employee training and development programs to enhance skills and ensure a positive customer experience.

4b) Calculate the Total & Average value of order price for each state.

Intuition:

Determine total and average order values for every state to understand regional sales performance.

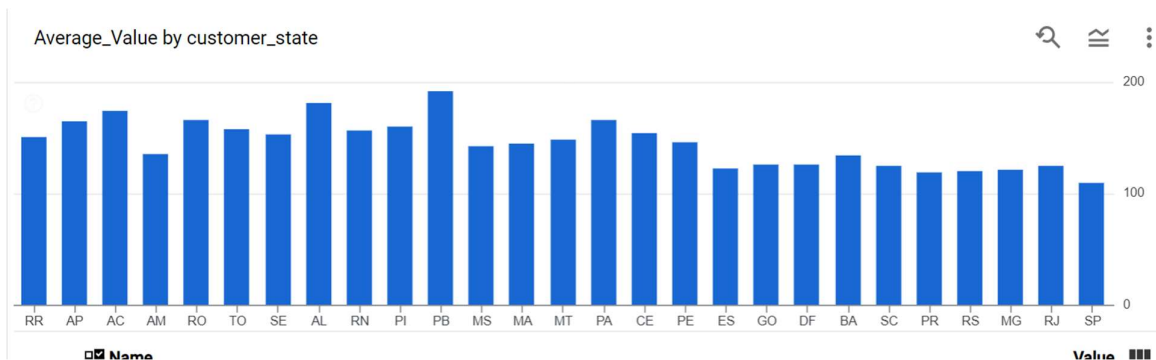
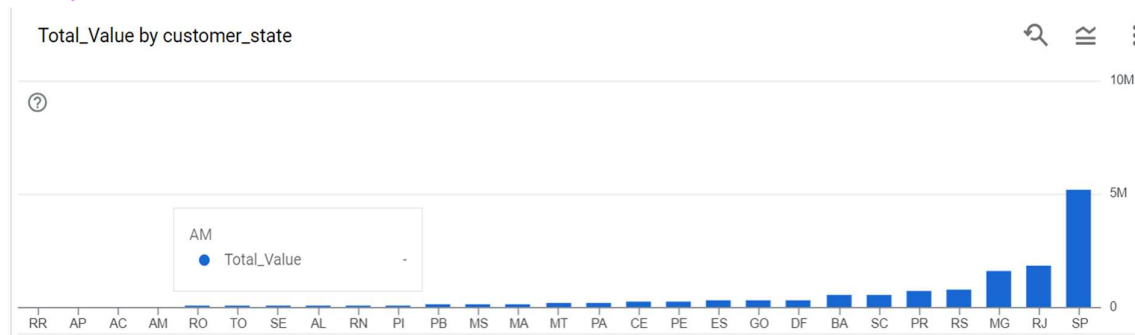
Query :

```
select
    c.customer_state,
    round(sum(oi.price),2) as Total_Value,
    round(avg(oi.price),2) as Average_Value
from `target_company_businesscase.customers` c
join `target_company_businesscase.orders` o
on c.customer_id = o.customer_id
join `target_company_businesscase.order_items` oi
on o.order_id = oi.order_id
group by c.customer_state
order by Total_value , Average_Value
limit 10
```

Result :

| Row | customer_state | Total_Value | Average_Value |
|-----|----------------|-------------|---------------|
| 1 | RR | 7829.43 | 150.57 |
| 2 | AP | 13474.3 | 164.32 |
| 3 | AC | 15982.95 | 173.73 |
| 4 | AM | 22356.84 | 135.5 |
| 5 | RO | 46140.64 | 165.97 |
| 6 | TO | 49621.74 | 157.53 |
| 7 | SE | 58920.85 | 153.04 |
| 8 | AL | 80314.81 | 180.89 |
| 9 | RN | 83034.98 | 156.97 |
| 10 | PI | 86914.08 | 160.36 |

Graph :



Insights :

- According to the graph, the total value of customers in São Paulo is higher compared to other states. This is because São Paulo is one of the most densely populated states in the country.
- Based on the graph of average values of states, the states with the highest average values are PB, AL, PA, RO, and RR.

Recommendations :

- Concentrate efforts on São Paulo, utilize tailored marketing to meet customer preferences and capitalize on higher population density advantage.
- Conduct a competitive analysis in both São Paulo and the states with high average values. Identify the key competitors in each region and analyse their strengths and weaknesses. This analysis will help you develop competitive advantages and position your brand effectively.
- Prioritize exceptional customer experience: timely service, easy purchase, post-purchase support. Satisfied customers become loyal advocates, boosting brand reputation and growth.
- Introduce region-specific products to meet unique demands, diversify offerings, attract more customers, and boost sales.

4c) Calculate the Total & Average value of order freight for each state.

Intuition:

Calculate total and average order freight values for each state to understand shipping costs, identify high-cost regions, and optimize logistics for better cost-efficiency.

Query:

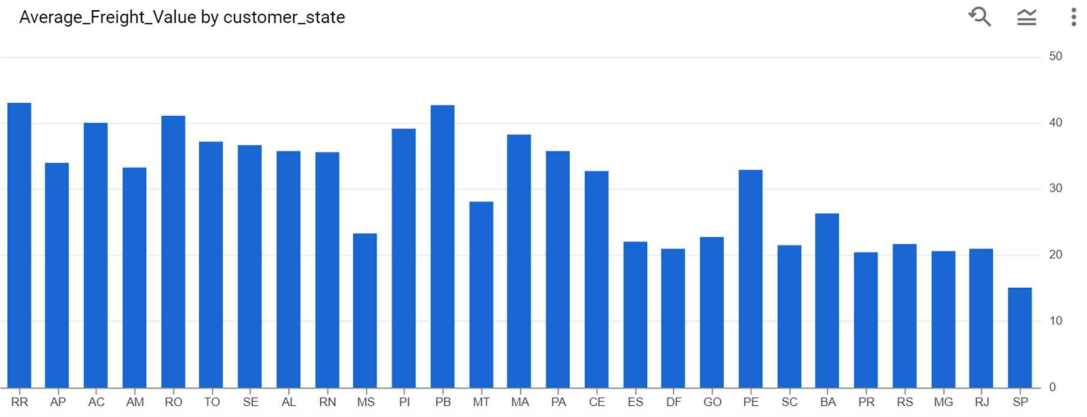
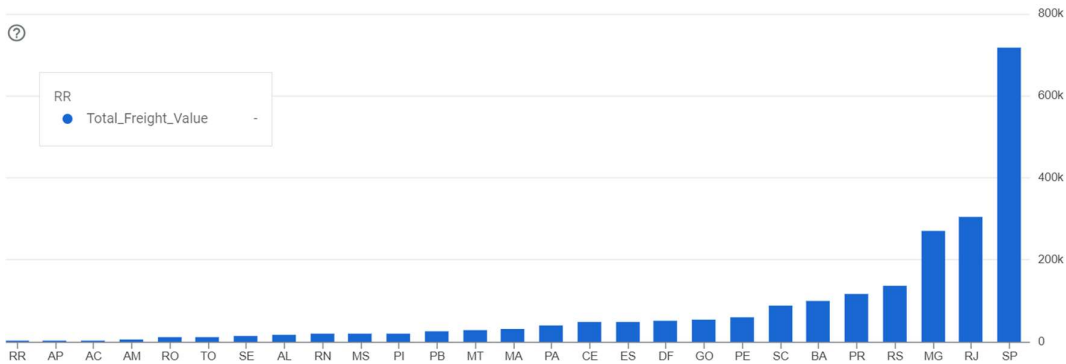
```
select
c.customer_state,
round(sum(oi.freight_value),2) as
Total_Freight_Value,
round(avg(oi.freight_value),2) as
Average_Freight_Value
from
`target_company_businesscase.customers`
c
join
`target_company_businesscase.orders` o
on c.customer_id = o.customer_id
join
`target_company_businesscase.order_items`
oi
on o.order_id = oi.order_id
```

```
group by c.customer_state
order by Total_Freight_Value,
Average_Freight_Value
limit 10
```

Result :

| Row | customer_state | Total_Freight_Value | Average_Freight_Val |
|-----|----------------|---------------------|---------------------|
| 1 | RR | 2235.19 | 42.98 |
| 2 | AP | 2788.5 | 34.01 |
| 3 | AC | 3686.75 | 40.07 |
| 4 | AM | 5478.89 | 33.21 |
| 5 | RO | 11417.38 | 41.07 |
| 6 | TO | 11732.68 | 37.25 |
| 7 | SE | 14111.47 | 36.65 |
| 8 | AL | 15914.59 | 35.84 |
| 9 | RN | 18860.1 | 35.65 |
| 10 | MS | 19144.03 | 23.37 |

Graph :



Insights :

- The freight_value of São Paulo is higher compared to other states.
- In terms of average freight_value, PB, PI, RO, and RR have higher values.
- The shipment cost for São Paulo is higher compared to other states in Brazil.

Recommendations :

- Investigate the reasons behind the higher freight_value in São Paulo. Explore potential ways to optimize shipping costs, negotiate better deals with logistics partners, or explore alternate shipping options to reduce expenses.
- Pay attention to PB, PI, RO, and RR with higher average freight_value. Research customer preferences and shipping needs. Negotiate better rates with carriers for São Paulo and high-cost states.
- Establish localized warehouses for strategic regions, educate customers on shipping costs and delivery times, and optimize packaging to reduce freight expenses.

5. Analysis based on sales, freight and delivery time

5a) Find the no. of days taken to deliver each order from the order's purchase date as delivery time. Also, calculate the difference (in days) between the estimated & actual delivery date of an order. Do this in a single query. You can calculate the delivery time and the difference between the estimated & actual delivery date using the given formula:

■ $\text{time_to_deliver} = \text{order_delivered_customer_date} - \text{order_purchase_timestamp}$

■ $\text{diff_estimated_delivery} = \text{order_estimated_delivery_date} - \text{order_delivered_customer_date}$

Intuition:

Using a single query, calculate delivery time by subtracting the order purchase date from the delivery date. Also, find the difference between the estimated delivery date and the actual delivery date for each order. This query allows efficient data retrieval and analysis.

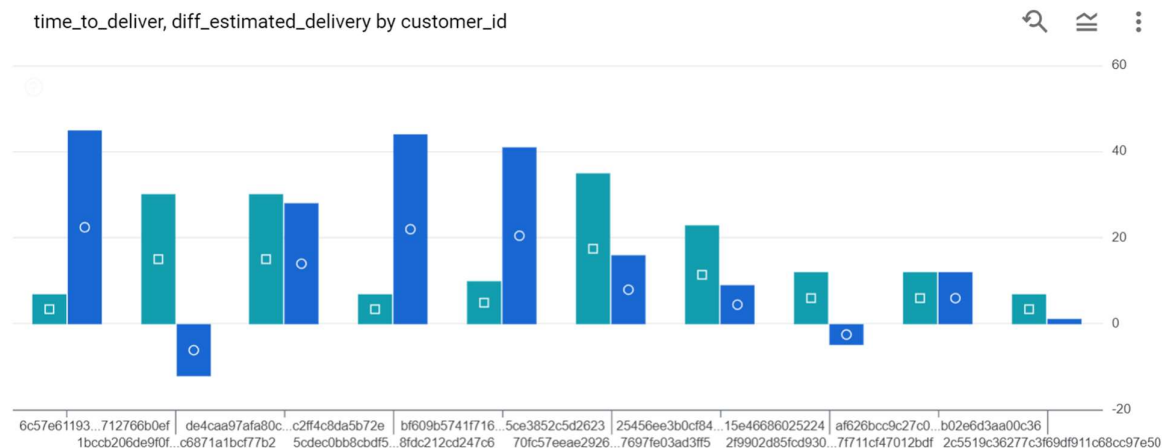
Query:

```
select
    order_id,
    order_purchase_timestamp,
    order_delivered_customer_date,
    order_estimated_delivery_date,
    date_diff(order_delivered_customer_date,
    order_purchase_timestamp, day) as
    time_to_deliver,
    date_diff(order_estimated_delivery_date,
    order_delivered_customer_date, day) as
    diff_estimated_delivery
from `target_company_businesscase.orders`
where order_delivered_customer_date is not
null
limit 10
```

Result:

| Row | order_id | order_purchase_timestamp | order_delivered_customer_date | order_estimated_delivery_date | time_to_deliver | diff_estimated_delivery |
|-----|-------------------------------|--------------------------|-------------------------------|-------------------------------|-----------------|-------------------------|
| 1 | 770d331c84e5b214bd9dc70a1... | 2016-10-07 14:52:30 UTC | 2016-10-14 15:07:11 UTC | 2016-11-29 00:00:00 UTC | 7 | 45 |
| 2 | 1950d777989f6a877539f5379... | 2018-02-19 19:48:52 UTC | 2018-03-21 22:03:51 UTC | 2018-03-09 00:00:00 UTC | 30 | -12 |
| 3 | 2c45c33d2f9cb8ff8b1c86cc28... | 2016-10-09 15:39:56 UTC | 2016-11-09 14:53:50 UTC | 2016-12-08 00:00:00 UTC | 30 | 28 |
| 4 | dabf2b0e35b423f94618bf965f... | 2016-10-09 00:56:52 UTC | 2016-10-16 14:36:59 UTC | 2016-11-30 00:00:00 UTC | 7 | 44 |
| 5 | 8beb59392e21af5eb9547ae1a... | 2016-10-08 20:17:50 UTC | 2016-10-19 18:47:43 UTC | 2016-11-30 00:00:00 UTC | 10 | 41 |
| 6 | 65d1e226dfaeb8cdc42f66542... | 2016-10-03 21:01:41 UTC | 2016-11-08 10:58:34 UTC | 2016-11-25 00:00:00 UTC | 35 | 16 |
| 7 | c158e9806f85a33877bdfd4f60... | 2017-04-14 22:06:32 UTC | 2017-05-08 11:10:26 UTC | 2017-05-18 00:00:00 UTC | 23 | 9 |
| 8 | b60b53ad0bb7dacacf2989fe2... | 2017-05-10 14:03:27 UTC | 2017-05-23 13:12:27 UTC | 2017-05-18 00:00:00 UTC | 12 | -5 |
| 9 | c830f223aae08493ebecb52f2... | 2017-04-22 15:50:30 UTC | 2017-05-05 13:27:50 UTC | 2017-05-18 00:00:00 UTC | 12 | 12 |
| 10 | a8aa2cd070eeac7e4368cae3d... | 2017-05-09 17:42:45 UTC | 2017-05-16 23:22:20 UTC | 2017-05-18 00:00:00 UTC | 7 | 1 |

Graph:



Insights:

- Order 1 received the fastest delivery within 7 days, although the estimated time was 45 days.
- Some orders exceeded the delivery time limit of 12 to 20 days, leading to delays in reaching customers.

Recommendations:

- Analyze the reasons for the significant difference between the actual delivery time for Orders and the estimated time. Identify areas for improvement in the delivery process to ensure faster and more accurate deliveries.
- Enhance communication with customers regarding estimated delivery times. Provide timely updates on the status of their orders and manage customer expectations to reduce dissatisfaction caused by delayed deliveries.

5b) Find out the top 5 states with the highest & lowest average freight value.

Intuition :

Identify the top 5 states with the highest and lowest average freight time. This analysis allows the company to focus on regions with efficient Freight Value.

Query :

```
with Average_Freight_values as (  
  select  
    c.customer_state,  
    round(avg(oi.freight_value),2)  
    Average_Freight_Value  
  from `target_company_businesscase.customers`  
  c  
  join `target_company_businesscase.orders` o  
  on c.customer_id = o.customer_id  
  join  
  `target_company_businesscase.order_items` oi  
  on o.order_id = oi.order_id  
  group by c.customer_state  
)  
  
CTE1 as (  
  select  
    customer_state,  
    Average_Freight_Value,  
    dense_rank() over(order by  
    Average_Freight_Value desc) as  
    Highest_Average_Freight_Value,  
    row_number() over(order by  
    Average_Freight_Value desc) as Row_No  
  from Average_Freight_values  
)  
  
CTE2 as (  
  select  
    customer_state,  
    Average_Freight_Value,  
    dense_rank() over(order by  
    Average_Freight_Value asc) as  
    Lowest_Average_Freight_Value,
```

```

row_number() over(order by
Average_Freight_Value asc) as Row_No
from Average_Freight_values
)

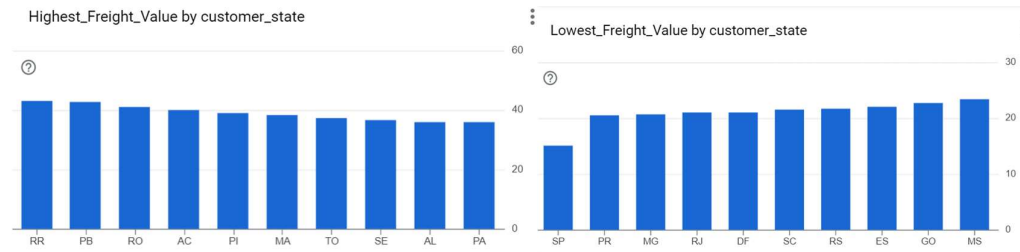
select
CTE1.customer_state as
Top_5Highest_average_FreightValue,
CTE1.Average_Freight_Value as
HighestFreightValue,
CTE2.customer_state as
Top_5_LowestFreightValue,
CTE2.Average_Freight_Value as
LowestFreightValue
from CTE1 inner join CTE2
on CTE1.Row_No = CTE2.Row_No
WHERE (CTE1.Highest_Average_Freight_Value
BETWEEN 1 AND 5) AND
(CTE2.Lowest_Average_Freight_Value BETWEEN 1
AND 5)
order by CTE1.Row_No

```

Result :

| Row | Top_5Highest_average_FreightValu | HighestFreightValue | Top_5_LowestFreightValue | LowestFreightValue |
|-----|----------------------------------|---------------------|--------------------------|--------------------|
| 1 | RR | 42.98 | SP | 15.15 |
| 2 | PB | 42.72 | PR | 20.53 |
| 3 | RO | 41.07 | MG | 20.63 |
| 4 | AC | 40.07 | RJ | 20.96 |
| 5 | PI | 39.15 | DF | 21.04 |

Graph :



Insights :

- According to the graph, the highest average freight_value is approximately the same for every state.
- According to the graph, the lowest average freight_value for every state is also approximately the same, but São Paulo has a very low freight value due to more distances to São Paulo.

Recommendations:

- Focus on regions with high demand and purchasing power due to similar average freight_values. Leverage São Paulo's significant customer base for growth by expanding operations and marketing efforts.
- Exceptional customer service for loyalty. Competitive pricing in states with similar freight_values, maintaining quality.

5c) Find out the top 5 states with the highest & lowest average delivery time.

Intuition :

Identify the top 5 states with the highest and lowest average delivery time. This analysis allows the company to focus on regions with efficient delivery and address issues in states with longer delivery times.

Query :

```
with CTE as (
  select
    c.customer_state,
```

```

        AVG(DATE_DIFF(o.order_delivered_customer_date,o.order_purchase_timestamp,DAY))
        Avg_Delivery_Time
    from
        `target_company_businesscase.orders` as
        o
    join
        `target_company_businesscase.customers`
        as c
    on c.customer_id = o.customer_id
    group by c.customer_state
),
CTE1 as (
    select
        customer_state,
        Avg_Delivery_Time,
        DENSE_RANK()OVER(ORDER BY
        Avg_Delivery_Time DESC)
        Highest_Delivery_Time,
        ROW_NUMBER()OVER(ORDER BY
        Avg_Delivery_Time DESC) ROW_NO
    from CTE
),
CTE2 as (
    select
        customer_state,
        Avg_Delivery_Time,
        DENSE_RANK()OVER(ORDER BY
        Avg_Delivery_Time asc)
        Lowest_Delivery_Time,
        ROW_NUMBER()OVER(ORDER BY
        Avg_Delivery_Time asc) ROW_NO
    from CTE
)

SELECT
    CTE1.customer_state AS
    Top5States_HighestAvgDeliveryTime,
    ROUND(CTE1.Avg_Delivery_Time,2) AS
    HighestAvgDeliveryTime,

```

```

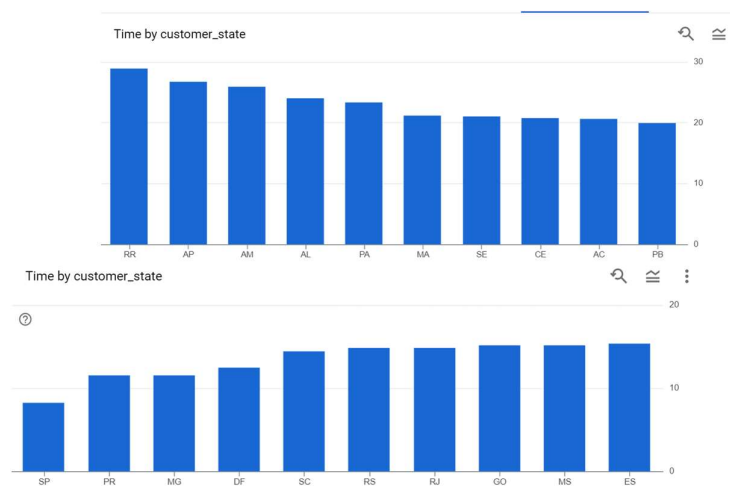
CTE2.customer_state AS
Top5States_Lowest_AvgDeliveryTime,
ROUND(CTE2.Avg_Delivery_Time,2) AS
LowestAvgDeliveryTime
FROM CTE1 INNER JOIN CTE2
ON CTE1.ROW_NO = CTE2.ROW_NO
WHERE (CTE1.Highest_Delivery_Time BETWEEN 1
AND 5) AND (CTE2.Lowest_Delivery_Time BETWEEN
1 AND 5)
ORDER BY CTE1.ROW_NO

```

Result :

| Row | Top5States_Highest_AvgDelivery | HighestAvgDeliveryTime | Top5States_Lowest_AvgDeliveryTime | LowestAvgDeliveryTime |
|-----|--------------------------------|------------------------|-----------------------------------|-----------------------|
| 1 | RR | 28.98 | SP | 8.3 |
| 2 | AP | 26.73 | PR | 11.53 |
| 3 | AM | 25.99 | MG | 11.54 |
| 4 | AL | 24.04 | DF | 12.51 |
| 5 | PA | 23.32 | SC | 14.48 |

Graph :



Insights :

- The states with the highest average delivery time in Brazil are RR, AP, and AM.
- The states with the lowest average delivery time in Brazil are SP, PR, and MG.

Recommendations :

- Analyze the reasons for the significant difference between the actual delivery time for Orders and the estimated time. Identify areas for improvement in the delivery process to ensure faster and more accurate deliveries.
- Enhance communication with customers regarding estimated delivery times. Provide timely updates on the status of their orders and manage customer expectations to reduce dissatisfaction caused by delayed deliveries

6) Analysis based on the payments:

6a). Find the month on month no. of orders placed using different payment types.

Intuition:

Analyze the number of orders placed each month, categorized by different payment types, to understand payment trends and identify any patterns or changes over time.

Query:

```
with CTE as (  
    SELECT  
    extract(month from  
    o.order_purchase_timestamp) as month,
```

```

extract(year from
o.order_purchase_timestamp) as year,
p.payment_type
FROM
`target_company_businesscase.payments` p
INNER
JOIN `target_company_businesscase.orders` o
ON p.order_id = o.order_id
)

```

```

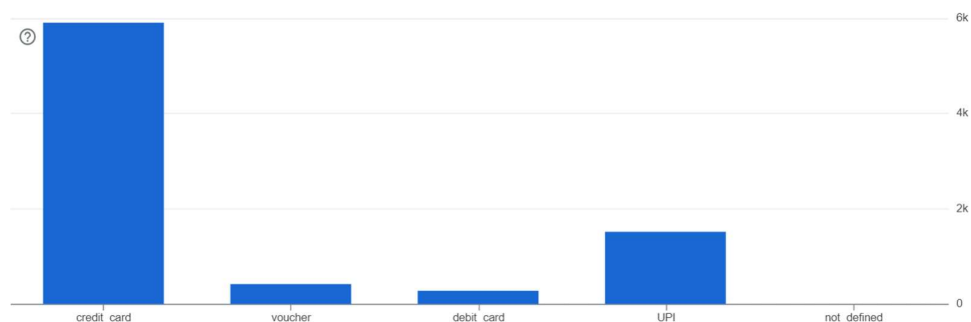
select
month,year,payment_type,
count(*) as Number_Of_Orders
from CTE
group by month,year,payment_type
order by year,CTE.month
limit 10

```

Result:

| Row | month | year | payment_type | Number_Of_Orders |
|-----|-------|------|--------------|------------------|
| 1 | 9 | 2016 | credit_card | 3 |
| 2 | 10 | 2016 | debit_card | 2 |
| 3 | 10 | 2016 | credit_card | 254 |
| 4 | 10 | 2016 | voucher | 23 |
| 5 | 10 | 2016 | UPI | 63 |
| 6 | 12 | 2016 | credit_card | 1 |
| 7 | 1 | 2017 | voucher | 61 |
| 8 | 1 | 2017 | UPI | 197 |
| 9 | 1 | 2017 | credit_card | 583 |
| 10 | 1 | 2017 | debit_card | 9 |

Graph :



Insights:

- According to the data, the highest payments were received via credit card.
- The next highest payment method is UPI (e.g., PhonePe, Google Pay).
- The least common payment method was "not_defined" (possibly cash on delivery or another undefined method).

Recommandations:

- Promote credit cards and UPI with discounts for secure, swift transactions. Ensure seamless processing, enhancing customer confidence. Streamline COD for a smooth experience.
- Expand digital payment options based on customer preferences for a broader audience. Gather feedback to understand preferences and pain points, enhancing the payment experience. Implement robust security measures to build trust in digital payment systems.

6b) Find the no. of orders placed on the basis of the payment instalments that have been paid.

Intuition:

Analyze orders based on paid instalments to understand the impact of payment options on purchase behaviour and preferences.

Query:

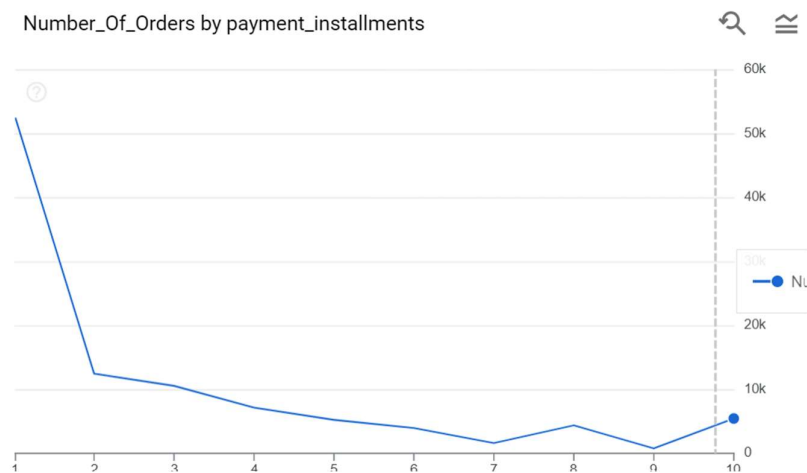
```
select
    payment_installments,
    count(order_id) as Number_Of_Orders
from `target_company_businesscase.payments`
where payment_installments > 0
group by payment_installments
```

```
order by payment_installments asc
limit 10
```

Result:

| Row | payment_installment | Number_Of_Orders |
|-----|---------------------|------------------|
| 1 | 1 | 52546 |
| 2 | 2 | 12413 |
| 3 | 3 | 10461 |
| 4 | 4 | 7098 |
| 5 | 5 | 5239 |
| 6 | 6 | 3920 |
| 7 | 7 | 1626 |
| 8 | 8 | 4268 |
| 9 | 9 | 644 |
| 10 | 10 | 5328 |

Graph:



Insights:

- According to analysis of data the number of orders decreasing when payment instalments increases

Recommendations:

- Provide flexible payment options for a broader audience, accommodating single-payment preferences. Incentivize full payment with discounts to increase order volumes. Transparently communicate total costs, including instalment-related charges and interests.
- If possible, provide customers with a range of instalment plans to choose from, catering to different financial needs.

-----X-----
