TARGET – BUSINESS CASE

Company Description :

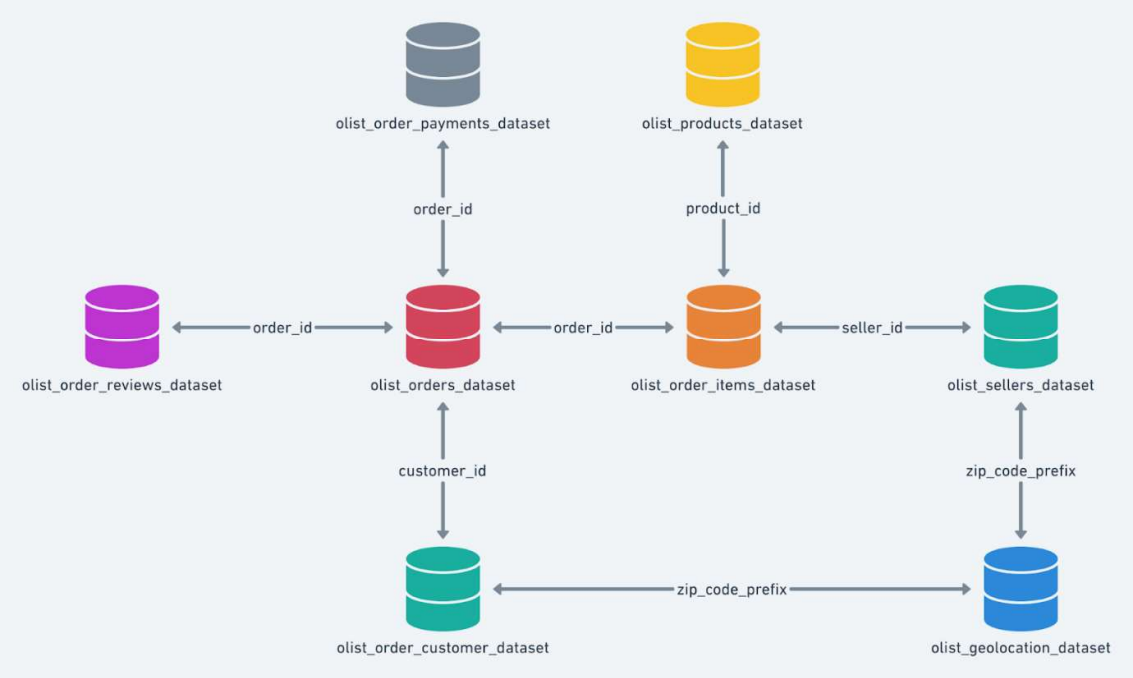
- Target is a globally renowned brand and a prominent retailer in the United States. Target makes itself a preferred shopping destination by offering outstanding value, inspiration, innovation, and an exceptional guest experience that no other retailer can deliver.

- This business case focuses on the operations of Target in Brazil and provides insightful information about 100,000 orders placed between Sep-2016 and Aug-2018. The dataset offers a comprehensive view of various dimensions including the order status, price, payment and freight performance, customer location, product attributes, and customer reviews.

- By analysing this extensive dataset, it becomes possible to gain valuable insights into Target's operations in Brazil. The information can shed light on various aspects of the business, such as order processing, pricing strategies, payment and shipping efficiency, customer demographics, product characteristics, and customer satisfaction levels.

Dataset:

Dataset schema:



The dataset consist of 8 tables such as , Customers Table, Geolocation Table, Order_items Table, Payments Table , Reviews Table, Orders Table, Products Table, Sellers Table …

1. checking the structure & characteristics of the dataset:

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

Query :

```
select column_name,data_type
from `target`.INFORMATION_SCHEMA.COLUMNS
where table_name= 'customers'
```

Output :

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 :

- We have queried to fetch specific information like the datatype of each column present in the 'customers' table. This helps for ensuring data integrity and can be helpful for understanding the table's structure, designing queries, or performing data analysis.

- "Customer ID" is typically a unique identifier assigned to each individual customer within a company's database system. It helps the company keep track of customer information, purchase history and other relevant data and this Id may be used for the tracking of order shipment and other requirements when the credentials of the customers are necessary.

- "Customer unique ID" refers to a more specific and individualized identifier assigned to a customer. This unique ID distinguishes one customer from the other in a more granular way. It may be used internally within Target's infrastructure for operational purposes, data analysis, or as a reference in their databases.

b. Time range between which the orders were placed.

Query :

```
select min(order_purchase_timestamp) as Min_time,
       max(order_purchase_timestamp) as Max_time
from `target.orders`
```

Output :

Row	Min_time	Max_time
1	2016-09-04 21:15:19 UTC	2018-10-17 17:30:18 UTC

Insights :

- We have queried to find the date and time of the First ever order placed and last order placed. Time period ranges approximately 2 Years for the given data.

- The time range is calculated to oversee the sales performance to identify trends, peak periods, optimize seasonal management, marketing strategies, promotions, analysing customer patterns, purchasing behaviour, order fulfilment rates, feedbacks, sales growth, customer reviews or customer support response times, annual events, can help identify seasonal trends and preferences. This information can guide and help in predicting future demand to improve overall customer satisfaction and to achieve operational Efficiency.

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

Query :

```
select count(distinct customer_city) as Cities,

count(distinct customer_state) as States

from `target.customers` c
join `target.orders` o
on c.customer_id = o.customer_id
```

Output :

Row	Cities	States
1	4119	27

Total Cities and States from the whole data

Query :

```
with cte as
(
select geolocation_city as City, geolocation_state as State
from `target.geolocation`
group by geolocation_city, geolocation_state

union all

select customer_city as City, customer_state as State
from `target.customers`
group by customer_city, customer_state

union all

select seller_city as City, seller_state as State
from `target.sellers`
group by seller_city, seller_state
)
select count(distinct City) as Total_Cities,
count(distinct State) as Total_States
from cte
```

Output :

Row	Total_Cities	Total_States
1	8126	27

Insights :

- Target has established a significant presence across various regions in Brazil, indicating a robust market position and a widespread customer base spanning the entire country. This highlights their extensive reach and strong foothold in the Brazilian market.

- We can infer that the count of unique no.of.cities and unique no.of.states in the geolocation has missed around 115 cities that are mentioned in union of whole dataset.

- The count of cities and states offers valuable information for conducting geographic analysis, identifying potential opportunities, customers based on locations and their behavioural pattern, evaluating regional performance, and facilitating decision-making processes to understand market potential and make informed decisions regarding resource allocation and strategic planning.

2. In-depth Exploration:

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

Query : (Month wise)

```
with cte as
(select concat(year,'-',mnth)as Year_Month,
        sum(no_of_orders) as Total_Orders,
        lag(sum(no_of_orders)) over(order by year,mnth)as lagg
 from (select extract(year from order_purchase_timestamp)as year,
             extract(month from order_purchase_timestamp)as mnth,
             count(*)as no_of_orders
      from `target.orders`
      where order_status not in ('canceled','unavailable')
      group by order_purchase_timestamp,year,mnth
      order by year, mnth)
group by year, mnth
order by year, mnth)
select Year_Month, total_orders, Growth_trend
From( select * ,
      concat(round(((total_orders-lagg)/lagg)*100,2),'  %') as
      Growth_trend
      from cte)
```

Output :

Row	Year_Month	total_orders	Growth_trend
1	2016-9	2	*null*
2	2016-10	293	14550 %
3	2016-12	1	-99.66 %
4	2017-1	787	78600 %
5	2017-2	1718	118.3 %
6	2017-3	2617	52.33 %
7	2017-4	2377	-9.17 %
8	2017-5	3640	53.13 %
9	2017-6	3205	-11.95 %
10	2017-7	3946	23.12 %
11	2017-8	4272	8.26 %
12	2017-9	4227	-1.05 %
13	2017-10	4547	7.57 %
14	2017-11	7423	63.25 %
15	2017-12	5620	-24.29 %

Query : (Year wise)

```
with cte as
(
select extract(year from order_purchase_timestamp) as Year,
       count(*) as No_of_Orders
from `target.orders`
where order_status<>'canceled'
group by 1
order by 1
)
select *, concat(round((No_of_Orders-YoY_orders)*100/YoY_orders,2), '%')
as Growth_Trend
from (select *, lag(No_of_Orders) over(order by Year) as YoY_orders
from cte
order by 1)
```

Output :

Row	Year	No_of_Orders	YoY_orders	Growth_Trend
1	2016	303	*null*	*null*
2	2017	44836	303	14697.36%
3	2018	53677	44836	19.72%

Data of 2016, 2017, 2018 is not fully available, that is why the No_of_Orders are less in 2016. Hence this data is not comparable year wise

Insights :

- By query analysis, the number of orders placed year-wise and month-wise excluding cancelled and unavailable orders, and considering the "Growth_trend" column indicating order growth rate compared to the previous year, it is evident that there was substantial order growth between 2016 and 2017, followed by a decrease in growth rate in 2018. This analysis helps identify patterns and trends in order placements over time.

- Overall, it demonstrates a consistent and significant upward trend in the number of orders in terms of orders placed whereas based on the revenue, there is a leap trend over the Brazil region, indicating positive business performance and increasing customer demand for Target's products in Brazil but must work on the revenue based to gain an uptrend is what we can from the chart analysis.

b. monthly seasonality in terms of the no. of orders being placed.

Query :

```
select Year,case
when Mnth=1 then 'January'
when Mnth=2 then 'Feburary'
when Mnth=3 then 'March'
when Mnth=4 then 'April'
when Mnth=5 then 'May'
when Mnth=6 then 'June'
when Mnth=7 then 'July'
when Mnth=8 then 'August'
when Mnth=9 then 'September'
when Mnth=10 then 'October'
when Mnth=11 then 'November'
else 'December'
end as Monthly_Season,
sum(No_of_Orders) as Total_Orders
from (select extract(year from order_purchase_timestamp) as Year,
       extract(month from order_purchase_timestamp) as Mnth,
       count(*) as No_of_Orders
from target.orders
where order_status <> 'canceled'
group by 1, 2
order by 1)
group by Year, Mnth
order by Year
```

Output :

Row	Year	Monthly_Season	Total_Orders
1	2016	October	300
2	2016	September	2
3	2016	December	1
4	2017	November	7507
5	2017	December	5662
6	2017	April	2386
7	2017	July	3998
8	2017	October	4605
9	2017	June	3229
10	2017	September	4265
11	2017	Feburary	1763
12	2017	January	797
13	2017	August	4304
14	2017	March	2649
15	2017	May	3671

Insights :

- The objective of the query was to calculate the total count orders for each month over the years to get an understanding and managing the business's order patterns.

- There is a clear monthly seasonal pattern in the number of orders, with increases and decreases throughout the year, perhaps influenced by holidays or specific shopping periods.

- We can infer that Peak is at November 2017 shows the highest number of orders, while September 2016 has the lowest number of orders. we can simply understand the uptrend in the orders place using chart analysis.

- Comparing with the world events and Brazilian culture, Having welcome the new year is great even , on the occasion they materialize the home appliances and needs, bath tubs, beds and hence the order were peaked during year end & November2017 is the month of Black Friday & they are more inclined towards soccer game and FIFA world cup 2018 happened , during that time orders peaked were placed. Therefore, we see the Monthly seasonality hotspots.

- Comparing the number of orders across different years reveals growth and trough in number of orders. There is significant growth from 2016 to 2017, but a decrease in average orders in 2018 compared to the previous year.

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

a. 0-6 hrs : Dawn
b. 7-12 hrs : Mornings
c. 13-18 hrs : Afternoon
d. 19-23 hrs : Night

Query :

```
select case
       when Hrs between 0 and 6 then 'Dawn'
       when Hrs between 7 and 12 then 'Morning'
       when Hrs between 13 and 18 then 'Afternoon'
       else 'Night'
       end as Order_Time,
       sum(No_of_Orders) as No_of_Orders_Placed
from(select extract(hour from order_purchase_timestamp) as Hrs,
     count(*) as No_of_Orders
from `target.orders`
group by 1)
group by 1
order by 2 desc
```

Output :

Row	Order_Time	No_of_Orders_Placed
1	Afternoon	38135
2	Night	28331
3	Morning	27733
4	Dawn	5242

Insights :

- We have queried to find the apt time as when the most of the orders are being placed and we have inferred the same using chart analysis.
- Based on the output, Brazilian customers tend to place the highest number of orders during the afternoon hours, followed by the night hours.
- Morning hours also show a substantial volume of orders, while dawn hours have fewer orders placed.
- This suggests that customers in Brazil are actively engaged in online shopping during the afternoon and evening, possibly during leisure time, with a significant number of orders placed in the morning as well.

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

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

Query :

```
select format_timestamp("%Y-%m", order_purchase_timestamp) as Year_Month,
       c.customer_state as Customer_State,
       count(*) as No_of_Orders_Placed
from `target.customers` c
join `target.orders` o
on c.customer_id = o.customer_id
group by 1,2
order by 1,2
```

Output:

Row	Year_Month	Customer_State	No_of_Orders_Placed
17	2016-10	PR	19
18	2016-10	RJ	56
19	2016-10	RN	4
20	2016-10	RR	1
21	2016-10	RS	24
22	2016-10	SC	11
23	2016-10	SE	3
24	2016-10	SP	113
25	2016-12	PR	1
26	2017-01	AC	2

Insights :

- The query states the number of orders placed, minimum and maximum order values, and state wise running average of orders placed. On a broader perspective, comparing order counts with minimum and maximum values helps identify states that consistently Peaks or troughs. This information allows prioritization of geolocations and targeted strategies to enhance orders in specific states.

- The state-wise average order counts indicate overall order count trends, highlighting states with steady or increasing order placed. This insight identifies growth opportunities and effective marketing strategies in different regions .

- In summary _ by analyzing this data, this enables data-driven decision making, aiding in production planning, considering seasonality, state-wise performance, and growth trends, to optimize marketing, operations, and resource allocation strategies across different states throughout the Brazil country

b. Distribution of customers across the states.

Query :

```
select customer_state as States,
       count(distinct customer_id) as No_of_Customers
from `target.customers`
group by 1
order by 2 desc
```

Output :

Row	States	No_of_Customers
1	SP	41746
2	RJ	12852
3	MG	11635
4	RS	5466
5	PR	5045
6	SC	3637
7	BA	3380
8	DF	2140
9	ES	2033
10	GO	2020

Insights :

- The query captures the total number of customers, number of customers whom made more than 1 purchase and total orders placed state wise helps us to do the Geographical Analysis.

- By examining customer distribution across states, regional trends and patterns can be identified. States like Sao Paulo (SP) and Rio de Janeiro (RJ) with a high count of customers indicate potential market hotspots. State with lower customer counts such as Roraima (RR) and Tocantins (TO) with only 45 customers, present potential hotspots for market opportunities for expansion and growth opportunities. States with a higher number of regular customers suggests reflects higher customer likability's or satisfaction.

- This information helps in decision making of where to concentrate marketing campaigns, expansion into new regions, improvement strategies for retaining customers, and identification of areas for growth and investment.

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

   a. Get the % increase in the cost of orders from year 2017 to 2018 (include months between Jan to Aug only).

Query :

```
with cte1 as
 (
  select * from `target.orders` o
  join `target.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
  ), cte2 as (
    select extract(year from order_purchase_timestamp) as Year,
    sum(payment_value) as cost
    from cte1
    group by Year
    order by Year
  )
  select *, (cost - lag(cost, 1) over(order by Year))*100 / lag(cost, 1)
over(order by Year) as perc_inc
  from cte2
```

Output:

Row	Year	cost	perc_inc
1	2017	3669022.119999...	*null*
2	2018	8694733.839999...	136.9768716466...

Insights :

- The query shows the percentage increase in the total cost month wise and year wise which signifies the increase in order costs of products suggests a pricing adjustment or changes in product offerings, influenced by factors like procurement, raw material costs, and demand-supply dynamics.

- The data shows a positive trend of increasing order amounts from 2017 to 2018, a substantial decrease in cost of the products makes us understand the economic growth of the country indicating potential growth and success in the business across Brazilian states.

- Understanding these market dynamics is to be considered to make the change and replicate successful strategies in the future. These insights help understand seasonal patterns and growth trends, supporting strategic planning and decision-making.

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

Query:(from payments table)

```
select State, Total_Payment, round(Total_Payment/Unique_Order_Id, 2) as
Avg_Payment
from(select c.customer_state as State, round(sum(p.payment_value), 2) as
Total_Payment,
count(distinct o.order_id) as Unique_Order_Id
from `target.customers` c
join `target.orders` o
on c.customer_id = o.customer_id
join `target.payments` p
on p.order_id = o.order_id
group by 1 order by 2 desc)
```

Output:

Row	State	Total_Payment	Avg_Payment
1	SP	5998226.96	143.69
2	RJ	2144379.69	166.85
3	MG	1872257.26	160.92
4	RS	890898.54	162.99
5	PR	811156.38	160.78
6	SC	623086.43	171.32
7	BA	616645.82	182.44
8	DF	355141.08	165.95
9	GO	350092.31	173.31
10	ES	325967.55	160.34

Query : (From order_items table)

```
select State, Total_Price, round(Total_Price/Unique_Order_Id, 2) as
Avg_Price
from(select c.customer_state as State, round(sum(oi.price), 2) as
Total_Price,
count(distinct o.order_id) as Unique_Order_Id
from `target.customers` c
join `target.orders` o
on c.customer_id = o.customer_id
join `target.order_items` oi
on oi.order_id = o.order_id
group by 1 order by 2 desc)
```

Output:

Row	State	Total_Price	Avg_Price
1	SP	5202955.05	125.75
2	RJ	1824092.67	142.93
3	MG	1585308.03	137.33
4	RS	750304.02	138.13
5	PR	683083.76	136.67
6	SC	520553.34	144.12
7	BA	511349.99	152.28
8	DF	302603.94	142.4
9	GO	294591.95	146.78
10	ES	275037.31	135.82

Insights :

- We have queried to find total and average amount of orders. This can be inferred from either payments table payment_value column or order_item tables price column.

- The analysis reveals significant variation in total and average order prices across different states. São Paulo (SP) and Rio de Janeiro (RJ) have higher total prices, indicating larger revenue contributions, while Roraima (RR) and Amapá (AP) have relatively lower total prices

- Amapá (AP) and Roraima (RR) have higher average order prices, suggesting potential higher-value customers or demand for premium-priced products, while states like Espírito Santo (ES) and Goiás (GO) have relatively lower average order prices.

- This information enables targeted marketing campaigns and provides insights into regional customer behavior and preferences.

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

Query :

```
select State, Total_Freight, round(Total_Freight/Unique_Order_Id, 2) as
Avg_Freight
from(select c.customer_state as State, round(sum(oi.freight_value), 2) as
Total_Freight,
count(distinct o.order_id) as Unique_Order_Id
from `target.customers` c
join `target.orders` o
on c.customer_id = o.customer_id
join `target.order_items` oi
on oi.order_id = o.order_id
group by 1 order by 2 desc)
```

Output :

Row	State	Total_Freight	Avg_Freight
1	SP	718723.07	17.37
2	RJ	305589.31	23.95
3	MG	270853.46	23.46
4	RS	135522.74	24.95
5	PR	117851.68	23.58
6	BA	100156.68	29.83
7	SC	89660.26	24.82
8	PE	59449.66	36.07
9	GO	53114.98	26.46
10	DF	50625.5	23.82

Insights :

- The query showcases the cost of freight values for each states and we infer that states like São Paulo (SP) and Minas Gerais (MG) have higher total freight values, indicating higher overall shipping costs incurred. On the other hand, states like Roraima (RR) and Amapá (AP) have lower total freight values.

- Average freight values reflect the average shipping cost per order, with states like São Paulo (SP) and Minas Gerais (MG) having lower average freight values, suggesting more efficient logistics networks or economies of scale.

- States like Paraíba (PB) and Piauí (PI) have higher average freight values, indicating relatively higher shipping costs per order. These insights can assist in optimizing logistics operations and understanding regional shipping trends.

5. **Analysis based on sales, freight and delivery time.**
   a. 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.

Query:

```
select order_id,
      datetime_diff(order_delivered_customer_date,
order_purchase_timestamp,day) as time_to_deliver,

     datetime_diff(order_delivered_customer_date,
order_estimated_delivery_date,day) as diff_estimated_delivery
from `target.orders`
```

Output:

Row	order_id	time_to_deliver	diff_estimated_delivery
1	1950d777989f6a877539f5379...	30	12
2	2c45c33d2f9cb8ff8b1c86cc28...	30	-28
3	65d1e226dfaeb8cdc42f66542...	35	-16
4	635c894d068ac37e6e03dc54e...	30	-1
5	3b97562c3aee8bdedcb5c2e45...	32	0
6	68f47f50f04c4cb6774570cfde...	29	-1
7	276e9ec344d3bf029ff83a161c...	43	4
8	54e1a3c2b97fb0809da548a59...	40	4
9	fd04fa4105ee8045f6a0139ca5...	37	1
10	302bb8109d097a9fc6e9cefc5...	33	5

Insights :

- From Query Analysis, The "Time_to_deliver" column represents the number of days taken to deliver an order to the customer from the purchase date, while the "diff_estimated_delivery" column indicates the difference between the estimated delivery date and the actual delivery date.

- By analyzing this data, we can identify orders that took longer to deliver and compare each delivery time with the average delivery timeline to assess delivery efficiency. Negative values in the "diff_estimated_delivery" column indicate delayed deliveries, while positive values indicate early deliveries.

- Digging deeper, the reasons for these variances can help improve delivery timelines and reduce the difference between estimated and actual delivery dates, leading to enhanced logistics and delivery processes.

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

Query:

```
select States, Avg_Freight
from
((select concat(State_Top_5_Highest, '    Top_5') as States,
round(Total_Freight/Unique_Order_Id, 2) as Avg_Freight
from(select c.customer_state as State_Top_5_Highest,
round(sum(oi.freight_value), 2) as Total_Freight,
count(distinct o.order_id) as Unique_Order_Id
from `target.customers` c
join `target.orders` o
on c.customer_id = o.customer_id
join `target.order_items` oi
on oi.order_id = o.order_id
group by 1)
order by 2 desc
limit 5)

union all

(select concat(State_Top_5_Highest, '    Low_5') as States,
round(Total_Freight/Unique_Order_Id, 2) as Avg_Freight
from(select c.customer_state as State_Top_5_Highest,
round(sum(oi.freight_value), 2) as Total_Freight,
count(distinct o.order_id) as Unique_Order_Id
from `target.customers` c
join `target.orders` o
on c.customer_id = o.customer_id
join `target.order_items` oi
on oi.order_id = o.order_id
group by 1)
order by 2 asc
limit 5)) as A
```

Output: (In single Table)

Row	States	Avg_Freight
1	RR Top_5	48.59
2	PB Top_5	48.35
3	RO Top_5	46.22
4	AC Top_5	45.52
5	PI Top_5	43.04
6	SP Low_5	17.37
7	MG Low_5	23.46
8	PR Low_5	23.58
9	DF Low_5	23.82
10	RJ Low_5	23.95

Insights :

- Our query shows that Certain states exhibit higher average freight values, indicating specific characteristics or logistical challenges that result in increased freight costs.

- Certain states exhibit higher average freight values, indicating specific characteristics or logistical challenges that result in increased freight costs. Some states demonstrate lower average freight values, suggesting more favorable logistics infrastructure or other factors contributing to reduced freight costs. With average freight costs, we can that there is a notable difference (Significant Variation) in the average freight value among different states.

- These insights provide businesses with valuable information for optimizing their shipping and logistics operations. It enables them to make informed decisions regarding pricing strategies, supply chain optimization, and resource allocation.

- For states with the lowest average freight values, such as São Paulo (SP), Distrito Federal (DF), Rio de Janeiro (RJ), Minas Gerais (MG), and Paraná (PR), businesses can focus on further optimizing freight costs by negotiating better rates, consolidating shipments, and improving logistics routes.

- On the other hand, states with the highest average freight values, such as Piauí (PI), Acre (AC), Rondônia (RO), Roraima (RR), and Paraíba (PB), can explore opportunities to mitigate factors contributing to higher costs, including partnering with local logistics providers, optimizing transportation routes, and leveraging technology for efficient freight management.

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

Query:

```
select State, Avg_del_time
from ((select concat(States, '       top_5_slow') as State,
round((time_to_deliver/Order_Id),2) as Avg_del_time
from(select count(distinct order_id) as Order_Id,
       c.customer_state as States,
       sum(datetime_diff(o.order_delivered_customer_date,
o.order_purchase_timestamp,day)) as time_to_deliver,
from `target.orders` o
join `target.customers` c
on c.customer_id = o.customer_id
group by 2)
order by 2 desc
limit 5)
union all
(select concat(States, '        top_5_Fast') as State,
round((time_to_deliver/Order_Id),2) as Avg_del_time
from(select count(distinct order_id) as Order_Id,
       c.customer_state as States,
       sum(datetime_diff(o.order_delivered_customer_date,
o.order_purchase_timestamp,day)) as time_to_deliver,
from `target.orders` o
join `target.customers` c
on c.customer_id = o.customer_id
group by 2)
order by 2 asc
limit 5)) as A
```

Output: (In single Table)

Row	State	Avg_del_time
1	SP top_5_Fast	8.05
2	PR top_5_Fast	11.25
3	MG top_5_Fast	11.27
4	DF top_5_Fast	12.16
5	SC top_5_Fast	14.12
6	AP top_5_slow	26.34
7	RR top_5_slow	25.83
8	AM top_5_slow	25.46
9	AL top_5_slow	23.11
10	PA top_5_slow	22.62

Insights :

- From our queried output, we can say that states with the fastest average delivery times, such as Alagoas (AL), Pará (PA), Amazonas (AM), Amapá (AP), and Roraima (RR), the logistics are perfectly fine whereas the states with slow deliveries such as Santa Catarina(SC), Minas Gerais (MG), Distrito Federal (DF), (PR), Sao Paulo(SP)are to be considered.

- It is essential to identify and address the factors causing these prolonged delivery durations. Businesses should evaluate their logistics network, transportation routes, and last-mile delivery processes to optimize efficiency. Collaborating with local logistics providers or establishing strategic partnerships can also help improve delivery performance in these states. Implementing technology solutions like real-time tracking systems and efficient delivery scheduling tools can streamline the last-mile operations.

- Open communication and collaborative relationships with shipping carriers and logistics partners are crucial for monitoring and addressing any issues promptly. By optimizing the last-mile delivery process and maintaining efficient relationships with partners, businesses can reduce delivery times, enhance customer experience, and improve overall satisfaction.

d. Find out the top 5 states where the order delivery is really fast as compared to the estimated date of delivery.
You can use the difference between the averages of actual & estimated delivery date to figure out how fast the delivery was for each state.

Query:

```
select States, round((estimated_del_time/Order_Id)-
(actual_del_time/Order_Id), 2) as delivery_time
from(select count(distinct order_id) as Order_Id,
       c.customer_state as States,
       sum(datetime_diff(o.order_delivered_customer_date,
o.order_purchase_timestamp,day)) as actual_del_time,
       sum(datetime_diff(o.order_estimated_delivery_date,
o.order_purchase_timestamp ,day)) as estimated_del_time
from `target.orders` o
join `target.customers` c
on c.customer_id = o.customer_id
where order_status = 'delivered'
group by 2)
order by 2 desc
limit 5
```

Output:

Row	States	delivery_time
1	AC	20.09
2	RO	19.47
3	AP	19.13
4	AM	18.94
5	RR	16.66

Insights :

- From the Query Analysis, the top 5 states like Amazonas (AM), Amapá (AP), Roraima (RR), Acre (AC), Rondônia(RO) have average speed for order delivery compared to the estimated date, with an average difference of 16 to 19 days. This suggests efficient delivery processes, well-established transportation networks, optimized routing strategies, and proactive coordination with shipping carriers.
- Customers benefit from reliable and prompt delivery, increasing satisfaction. Businesses can leverage this information to highlight their effective logistics operations and gain a competitive advantage. It also helps identify areas for improving delivery efficiency by comparing with states that have longer delivery times.

6. Analysis based on the payments:

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

Query :

```
select extract(year from order_purchase_timestamp) as Year,
extract(month from order_purchase_timestamp) as Month,
payment_type,
count(distinct o.order_id) as No_of_Orders
from `target.payments` p
join `target.orders` o
on o.order_id = p.order_id
group by 1, 2, 3
order by 1 asc, 2 asc, 4 desc
```

Output :

Row	Year	Month	payment_type	No_of_Orders
1	2016	9	credit_card	3
2	2016	10	credit_card	253
3	2016	10	UPI	63
4	2016	10	voucher	11
5	2016	10	debit_card	2
6	2016	12	credit_card	1
7	2017	1	credit_card	582
8	2017	1	UPI	197
9	2017	1	voucher	33
10	2017	1	debit_card	9

Insights :

- The query provides insights into customer payment preferences by analyzing the monthly distribution of payment types. By analyzing the monthly distribution of payment types, businesses can gain insights into customer payment preferences, identify trends in payment methods, This information can be valuable for optimizing the checkout experience, expanding payment options and improving overall customer satisfaction.

- Credit card emerges as the preferred payment method, highlighting its convenience, security, and widespread acceptance. Businesses should ensure seamless credit card payment processing and maintain strong partnerships with payment service providers. Analyzing payment trends helps optimize the checkout experience, expand payment options, and enhance overall customer satisfaction.

b. Find the no. of orders placed on the basis of the payment installments that have been paid.

Query :

```
select payment_installments,
count(distinct order_id) as No_of_Orders
from `target.payments`
where payment_installments <> 0
group by 1
order by 1
```

Output :

Row	payment_installment	No_of_Orders
1	1	49060
2	2	12389
3	3	10443
4	4	7088
5	5	5234
6	6	3916
7	7	1623
8	8	4253
9	9	644
10	10	5315

Insights :

- Our Query Results shows the count of distinct orders for each payment instalment option. It indicates the distribution of customers who choose different instalment plans for their payments. Mostly preferred instalment is 1 or we can say that atleast 1 of many has been paid. The data indicates a notable customer preference for a particular focus on the 9-10 instalment range.

- To leverage this pattern and drive sales growth, businesses should consider providing customers with flexible payment options and instalment plans. By offering the ability to divide payments into multiple instalments, businesses can accommodate customer preferences, enhance affordability, and potentially expand their customer base. Incorporating instalment plans into the payment options can be an effective strategy to boost sales and improve overall customer satisfaction.

Recommendations :

- The count current customers are only from 4113 cities whereas Brazil has over 8126 and 27 states ,In order to give accessibility to the mass customers, TARGET should target even more on expanding across the Brazilian country.

- To address low-order months, businesses should conduct market research, collaborate with complementary businesses, offer promotions, and utilize targeted marketing. During different times of the day, strategies should be tailored to maximize sales. In states with higher customer bases, prioritize customer engagement, while in states with lower customer bases, target growth opportunities. Identify and leverage competitive advantages to stand out in the market

- Also inferring favourite product category, campaigning & marketing seasonal products related to the Brazilian culture like NEWYEAR, BLACK FRIDAY, CARNIVALS, FIFA WORLD CUP & CAPOEIRA helps us in understanding their which in turn gains more audience.

- To improve businesses, Optimize shipping operations, pricing strategies, and resource allocation. Focus on reducing the costs through negotiations and route improvements by having partnerships with local carriers and logistics partners are vital so that businesses can reduce delivery times and enhance logistics processes by using technology solutions which helps us improve performance by identifying the causing factors as last-mile delivery processes can streamline operations and reduce transit times. So that we can ensure customer satisfaction.

- Maintain proactive communication with customers about delivery expectations and provide timely updates on order status and potential delays.

- Ensure a secure payment infrastructure that supports various payment methods, adapting to evolving customer preferences and emerging technologies and also educate customers about alternative payment methods and offer incentives to encourage their usage. Promote the benefits of lower installment options and create value through targeted marketing campaigns.

- Collect customer feedback and analyze behaviour to understand preferences and to ensure customer satisfaction. With that , Branding can be gained.