TITLE: Business Case Study – TARGET SQL

Context:

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 2016 and 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.

Problem Statement:

Assuming you are a data analyst/ scientist at Target, you have been assigned the task of analysing the given dataset to extract valuable insights and provide actionable recommendations.

What does 'good' look like?

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

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

    Query:

    ```
    SELECT COLUMN_NAME, DATA_TYPE
    FROM `Project1TargetSQL.INFORMATION_SCHEMA.COLUMNS`
    WHERE TABLE_NAME = 'customers'
    ```

    Output screenshot:

    Query results

    SAVE RESULTS EXPLORE DATA

    JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

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

    <u>Insights</u>:

All columns in this example are saved as strings (VARCHAR), except customer_zip_code_prefix which is integer. This implies that the data might mostly be textual in nature, and that any values relating to dates or numbers are probably saved as strings

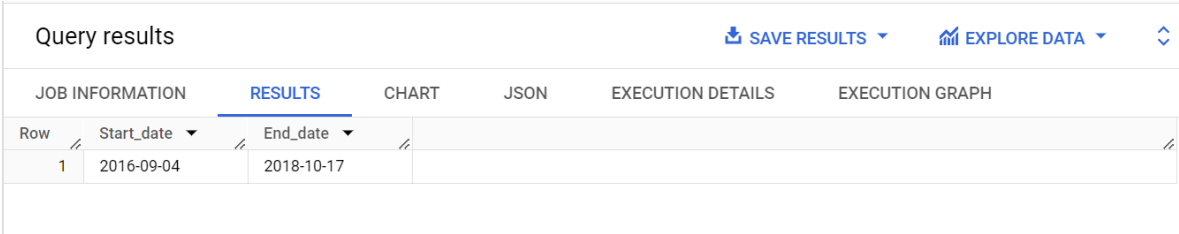
Recommendations: NA

2. Get the time range between which the orders were placed.

Query:

```
SELECT min(dt) as Start_date,
        max(dt) as End_date
FROM
  (
        select extract(date from order_purchase_timestamp) as dt
        from  `Project1TargetSQL.orders`
  ) t
```

Output screenshot:



Row	Start_date	End_date
1	2016-09-04	2018-10-17

Insights:

1. The orders in this case were placed between 2016-09-04 and 2018-10-17.

2. To analyze trends, seasonality, and overall order patterns over a certain time, it can be helpful to know the time range of the orders.

Recommendations: NA

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

Query:

```
SELECT
```

```
        count(distinct geolocation_city) as city_count ,count(distinct
    geolocation_state) as state_count
FROM
    `Project1TargetSQL.geolocation`
```

Output screenshot:

Query results SAVE RESULTS EXPLORE DATA

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	city_count	state_count
1	8011	27

<u>Insights</u>:

1. The dataset in this example has **27** distinct states and **8011** distinct cities. These figures can aid in our comprehension of the geographic distribution of our clientele or the scope of our dataset.

2. Analysing the distribution of cities and states might reveal information about the diversity or concentration of our clientele in various geographic areas. It might be helpful for regional analysis, figuring out hotspots, or figuring out how far our company has spread across the nation or the globe.

<u>Recommendations</u>: NA

2. In-depth Exploration:

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

Query:

```
select
    extract(year from order_purchase_timestamp) as Year,
     count(*) as number_of_orders_per_year
from
    `Project1TargetSQL.orders`
group by Year
order by 1
```

Output screenshot:

Query results

SAVE RESULTS EXPLORE DATA

JOB INFORMATION RESULTS CHART JSON EXECUTION DETAILS EXECUTION GRAPH

Row	Year	number_of_orders_per_year
1	2016	329
2	2017	45101
3	2018	54011

Insights:

1. There has been an upward trend in the number of orders over the past few years after examining the results. A favourable trend can be seen if the order number regularly rises year over year. If there are variations or a negative tendency, however, it points to a different pattern.

Recommendations:

1.With an increasing number of orders, it's crucial to optimize inventory management. Ensure that the inventory levels are sufficient to meet the growing demand while avoiding overstocking.
2. Use the historical order data to improve forecasting and planning.

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

Query

```
select
    extract(month from order_purchase_timestamp) as Month_wise_purchase,
    count(order_id) as number_of_orders
from
`Project1TargetSQL.orders`
group by Month_wise_purchase
order by 1
```

Output screenshot:

Query results

SAVE RESULTS EXPLORE DATA

JOB INFORMATION RESULTS CHART JSON EXECUTION DETAILS EXECUTION GRAPH

Row	Month_wise_purchas	number_of_orders
1	1	8069
2	2	8508
3	3	9893
4	4	9343
5	5	10573
6	6	9412
7	7	10318
8	8	10843
9	9	4305
10	10	4959
11	11	7544
12	12	5674

Insights:

We see high purchase during May, June, July and August and a decline during October to December.

Recommendations:

1.Plan and execute targeted marketing campaigns during the peak months (May to August) to capitalize on the high purchase behaviour.
2.Tailor promotions, discounts, or special offers to attract and retain customers during these months.
3. During the months of lower purchasing activity (October to December), implement targeted promotions to stimulate sales.
4.Consider clearance sales, holiday promotions, or other incentives to maintain customer interest.

3. During what time of the day, do the Brazilian customers mostly place their orders? (Dawn, Morning, Afternoon or Night)
    1. 0-6 hrs : Dawn
    2. 7-12 hrs : Mornings
    3. 13-18 hrs : Afternoon
    4. 19-23 hrs : Night

Query:

```
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 'Morning'

    WHEN EXTRACT(HOUR FROM order_purchase_timestamp) BETWEEN 13 AND 18 THEN 'Afternoon'
```

```
        WHEN EXTRACT(HOUR FROM order_purchase_timestamp) BETWEEN 19
        AND 23 THEN 'Night'

        END AS order_time_interval,

COUNT(*) AS order_count


FROM `Project1TargetSQL.orders`

GROUP BY order_time_interval

ORDER BY order_count DESC
```

Output screenshot:

Query results

SAVE RESULTS EXPLORE DATA

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	order_time_interval	order_count
1	Afternoon	38135
2	Night	28331
3	Morning	27733
4	Dawn	5242

<u>Insights</u>:

1.The highest order count is observed during the "Afternoon" time interval, indicating that this period is the peak time for placing orders.

2.There is a significant difference in order counts between the highest (Afternoon) and lowest (Dawn) time intervals, suggesting variability in customer ordering behaviour throughout the day

3.The "Night" and "Morning" time intervals have relatively high order counts, although not as high as the "Afternoon." This may indicate that there is substantial ordering activity during both early morning and late-night hours

<u>Recommendations</u>:

1.Businesses may consider strategizing promotions, marketing, or order fulfillment processes based on the observed time intervals. For instance, during the peak afternoon period, ensuring efficient order processing and customer service may be crucial.

2.The insights can be used to make operational adjustments, such as optimizing staffing levels, inventory management, and marketing efforts to align with the observed ordering patterns.

3.Implementing time-sensitive marketing strategies, such as flash sales or exclusive offers during specific time intervals, could capitalize on the observed ordering trends.

4.Enhance the user experience during peak ordering times, ensuring that website/app performance, customer support, and order processing are optimized.

3. **Evolution of E-commerce orders in the Brazil region:**
   1. Get the month-on-month no. of orders placed in each state.

Query:

```
SELECT
        EXTRACT(MONTH  FROM  o.order_purchase_timestamp) AS order_month,
        c.customer_state,
        COUNT(*) AS number_of_order
FROM
        `Project1TargetSQL.orders` AS o
JOIN
        `Project1TargetSQL.customers` AS c
ON
        o.customer_id = c.customer_id
GROUP BY
        order_month,
        c.customer_state
ORDER BY
        order_month,
        c.customer_state
```

Output screenshot:

Query results

SAVE RESULTS EXPLORE DATA

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	order_month	customer_state	number_of_order
1	1	AC	8
2	1	AL	39
3	1	AM	12
4	1	AP	11
5	1	BA	264
6	1	CE	99
7	1	DF	151
8	1	ES	159
9	1	GO	164
10	1	MA	66

Results per page: 50 1 – 50 of 322

Same output ordered by number_of_orders in desc :

Query results

SAVE RESULTS EXPLORE DATA

JOB INFORMATION RESULTS CHART JSON EXECUTION DETAILS EXECUTION GRAPH

Row	order_month	customer_state	number_of_order
1	8	SP	4982
2	5	SP	4632
3	7	SP	4381
4	6	SP	4104
5	3	SP	4047
6	4	SP	3967
7	2	SP	3357
8	1	SP	3351
9	11	SP	3012
10	12	SP	2357

Results per page: 50 1 – 50 of 322

Insights:

1. We can learn more about the monthly order count for each state by examining the query's results. Over time, we can spot trends, patterns, or seasonality in the order volume for various states. We can use it to determine which states have consistently high order volumes and to pinpoint any months or states where order counts have significantly changed. Here in our data, we can find that for every month the state called **SP** has the **highest** number of orders.
2. States with lower population density, such as AC and AP have relatively lower order counts

Recommendations:

1.Concentrate marketing efforts and resources on states with consistently high order counts, such as SP and RJ. Implement targeted campaigns to maintain and potentially expand market share in these regions.
2.Explore opportunities for market penetration in less populated states, such as AC and AP. Develop strategies to increase brand awareness, customer engagement, and market share in these regions.

2. How are the customers distributed across all the states?

Query:

```
select
        customer_state,
        count(distinct customer_id) as number_of_customer
from
        `Project1TargetSQL.customers`
group by customer_state
order by number_of_customer desc
```

Output screenshot:

Query results

SAVE RESULTS EXPLORE DATA

JOB INFORMATION RESULTS CHART JSON EXECUTION DETAILS EXECUTION GRAPH

Row	customer_state	number_of_custome
1	SP	41746
2	RJ	12852
3	MG	11635
4	RS	5466
5	PR	5045
6	SC	3637
7	BA	3380
8	DF	2140

Load more

Results per page: 50 1 – 27 of 27

<u>Insights</u>:

1.SP,RJ, and MG have the highest numbers of customers, indicating a concentrated market presence in these states.Businesses should strategically focus on these states for marketing initiatives, customer engagement, and product/service optimization.

2.States like RS, PR, SC, and BA have a moderate number of customers.

Businesses may find growth opportunities by targeting these states with tailored marketing strategies.

3.For states with a lower number of customers, such as AC, AP, RR, and others, businesses should consider implementing targeted strategies to increase market share and customer engagement.

<u>Recommendations</u>:

1.Prioritize states with higher customer counts (SP, RJ, MG) for more significant investments.
2.Tailor marketing strategies to cater to the unique characteristics and preferences of customers in each state. Consider regional nuances, cultural factors, and local trends when designing campaigns.
3.Develop targeted market expansion plans for states with lower customer counts. Explore untapped markets by conducting market research and understanding the specific needs of customers in these regions.
4.Explore partnerships with local businesses or influencers in each state. Collaborations can help increase brand visibility and credibility within specific regions.

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

    1. 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.

Query:

```
with cte as (
        select
                format_datetime('%Y-%m',order_purchase_timestamp) as
                date_month_order,
                p.order_id,
                p.payment_value
        from
                `Project1TargetSQL.payments` p
        Join
                `Project1TargetSQL.orders` o
        on
                p.order_id = o.order_id
),
final as (
        select
                sum (
                case when date_month_order between '2017-01' and '2017-08'
                then payment_value
                end   ) as total_pay_2017,
                sum (
                case when date_month_order between '2018-01' and '2018-08'
                then payment_value
                end ) as total_pay_2018
                from cte
        )
        select
        round(((total_pay_2018-total_pay_2017)/total_pay_2017 * 100),2) as
        percentage_increase
        from final
```

Output screenshot:

Query results

SAVE RESULTS EXPLORE DATA

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	percentage_increase
1	136.98

Insights:

1. For both 2017 and 2018, only orders placed from January to August are considered.

2. To get the % increase, the query analyses the monthly prices between 2017 and 2018.

3. The findings tell us a growth rate of approximately 137% from 2017 to 2018.

Recommendations:

1.Conduct a detailed analysis to identify the factors contributing to the observed growth rate. Explore whether the increase is driven by higher order volumes, increased prices, or a combination of both.

2.Perform a product-level analysis to understand which products experienced the most significant price increases or contributed the most to the overall growth. This insight can guide inventory management and pricing strategies.

3.Capitalize on the significant growth rate of approximately 137% from 2017 to 2018. Use this positive trend to inform marketing, sales, and business development strategies, showcasing the company's success and attracting stakeholders.

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

Query:

```
SELECT
        customer_state,
        ROUND(SUM(p.payment_value),2) AS total_order_price,
        ROUND(AVG(p.payment_value),2) AS average_order_price
FROM
        `Project1TargetSQL.payments` p
JOIN
        `Project1TargetSQL.orders`  o
ON
        p.order_id = o.order_id
JOIN
```

```
        `Project1TargetSQL.customers`  c
ON
        o.customer_id = c.customer_id
GROUP BY
        customer_state
ORDER BY
        total_order_price DESC
```

Output screenshot:

Query results

SAVE RESULTS EXPLORE DATA

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	customer_state	total_order_price	average_order_price
1	SP	5998226.96	137.5
2	RJ	2144379.69	158.53
3	MG	1872257.26	154.71
4	RS	890898.54	157.18
5	PR	811156.38	154.15
6	SC	623086.43	165.98
7	BA	616645.82	170.82
8	DF	355141.08	161.13
9	GO	350092.31	165.76
10	ES	325967.55	154.71

Results per page: 50 1 – 27 of 27

<u>Insights</u>:

1.SP has the highest total order price, indicating a significant contribution to overall revenue. Businesses should continue to focus on optimizing operations and marketing efforts in this high-revenue state.

2.While SP has the highest total order price, PB has the highest average order price. Analyze the factors contributing to higher average order prices in specific states to understand customer behavior and preferences.

3.States like MA and RR show potential for growth, with relatively lower total order prices. Explore strategies to increase market share in these states, such as targeted marketing campaigns or promotions.

<u>Recommendations</u>:

1.Allocate resources and marketing efforts to states with both high total order prices and high average order prices, such as SP and PB.
2.Analyze and optimize pricing strategies in states with lower average order prices to enhance overall revenue. Consider competitive pricing and promotional adjustments.
3.Develop regionalized marketing campaigns tailored to each state's preferences, focusing on customer segmentation and enhancing the overall customer experience.

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

Query:

```
SELECT
        customer_state,
        ROUND(SUM(oi.freight_value),2) AS total_freight_value,
        ROUND(AVG(oi.freight_value),2) AS average_freight_value
FROM
        `Project1TargetSQL.order_items`  oi
JOIN
        `Project1TargetSQL.orders`  o
ON
        oi.order_id = o.order_id
JOIN
        `Project1TargetSQL.customers`  c
ON
        o.customer_id = c.customer_id
GROUP BY
        customer_state
ORDER BY
        total_freight_value DESC
```

Output screenshot:

Query results

SAVE RESULTS EXPLORE DATA

JOB INFORMATION RESULTS CHART JSON EXECUTION DETAILS EXECUTION GRAPH

Row	customer_state	total_freight_value	average_freight_valu
1	SP	718723.07	15.15
2	RJ	305589.31	20.96
3	MG	270853.46	20.63
4	RS	135522.74	21.74
5	PR	117851.68	20.53
6	BA	100156.68	26.36
7	SC	89660.26	21.47
8	PE	59449.66	32.92
9	GO	53114.98	22.77
10	DF	50625.5	21.04

Results per page: 50 1 – 27 of 27

Insights:

1.SP incurs the highest total freight cost, suggesting a significant portion of shipping expenses originates from this state.

2.Average freight costs vary among states, with PB having the highest average. Understanding regional differences can inform logistics and pricing strategies.

Recommendations:

1.Consider strategic pricing adjustments in states with higher average freight costs to maintain competitiveness and balance overall cost-effectiveness.
2.Explore ways to optimize logistics and shipping processes in states with higher freight costs, such as MA and RR, to potentially reduce overall expenses.

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

1. 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:
   1. **time_to_deliver** = order_delivered_customer_date - order_purchase_timestamp
   2. **diff_estimated_delivery** = order_delivered_customer_date - order_estimated_delivery_date

Query:

```
select
    order_id,
```

```
    date_diff(date(order_delivered_customer_date),
date(order_purchase_timestamp),day) as  time_to_deliver,

    date_diff(date(order_estimated_delivery_date),  date(order_delivered_custo
mer_date),day) as diff_estimated_delivery

from

    Project1TargetSQL.orders

 where

    order_status = 'delivered' AND order_delivered_customer_date IS NOT NULL

order by

   diff_estimated_delivery desc
```

Output screenshot:

Row	order_id	time_to_deliver	diff_estimated_deliv
1	0607f0efea4b566f1eb8f7d3c2...	3	147
2	c72727d29cde4cf870d569bf6...	7	140
3	eec7f369423b033e549c02f3c...	21	135
4	c2bb89b5c1dd978d507284be...	17	124
5	40dc2ba6f322a17626aac6244...	8	109
6	1a695d543b7302aa9446c8d5f...	12	84
7	39e0115911bf404857e14baa7...	12	83
8	38930f76efb00b138f4d632e4d...	11	78
9	c5132855100a12d63ed4e8ae0...	12	78
10	559eea5a72341a4c82dbce988...	13	78

Results per page: 50 1 – 50 of 96470

<u>Insights</u>:

1. Insights into the effectiveness of the delivery process, including any delays or early deliveries compared to the projected timeframe, can be gained by analyzing the delivery_time and diff_estimated_delivery columns.

2. These columns can be further examined to find trends, outliers, or elements that affect delivery times or discrepancies between estimated and actual delivery dates.

3. These insights can be applied to manage customer expectations, enhance customer satisfaction, optimize the delivery process, and improve logistics operations.

<u>Recommendations</u>:

1.Conduct a detailed analysis of the delivery_time and diff_estimated_delivery columns to assess the overall effectiveness of the delivery process. Identify patterns, delays, or instances of early deliveries to gain comprehensive insights into the delivery timeline.
2.Apply the gained insights to manage customer expectations effectively. Communicate realistic delivery timelines based on the analysis, reducing the likelihood of discrepancies between estimated and actual delivery dates and enhancing overall customer satisfaction.

3.Establish a continuous monitoring system to track delivery performance over time. Regularly revisit the analysis to adapt strategies as needed, ensuring ongoing improvement and alignment with changing customer expectations and market dynamics.

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

Query:

```
WITH
 high_value_state AS (
 SELECT
        customer_state AS high_avg_freight_value_state,
        ROUND(AVG(oi.freight_value),2) AS hight_average_freight_value,
        ROW_NUMBER() OVER (ORDER BY ROUND(AVG(oi.freight_value),2) DESC) as
rw_numb1
 FROM
         `Project1TargetSQL.order_items`  oi
 JOIN
        `Project1TargetSQL.orders`  o
 ON
        oi.order_id = o.order_id
 JOIN
        `Project1TargetSQL.customers` c
 ON
        o.customer_id = c.customer_id
 GROUP BY
        customer_state
 ORDER BY
        hight_average_freight_value DESC
 LIMIT    5 ),
 low_value_state AS (
 SELECT
```

```
        customer_state AS low_avg_freight_value_state,
         ROUND(AVG(oi.freight_value),2) AS low_average_freight_value,
         ROW_NUMBER() OVER (ORDER BY ROUND(AVG(oi.freight_value),2) ASC) as
rw_numb2
 FROM
        `Project1TargetSQL.order_items`  oi
 JOIN
        `Project1TargetSQL.orders`  o
 ON
        oi.order_id = o.order_id
 JOIN
        `Project1TargetSQL.customers`  c
 ON
         o.customer_id = c.customer_id
 GROUP BY
        customer_state
 ORDER BY
        low_average_freight_value ASC
 LIMIT  5 )
SELECT
         high_avg_freight_value_state,
        hight_average_freight_value,
        low_avg_freight_value_state,
        low_average_freight_value
FROM
        high_value_state hvs
JOIN
        low_value_state lvs
ON
        hvs.rw_numb1 = lvs.rw_numb2
```

Output screenshot:

Query results

SAVE RESULTS EXPLORE DATA

JOB INFORMATION RESULTS CHART JSON EXECUTION DETAILS EXECUTION GRAPH

Row	high_avg_freight_value_state	hight_average_freigh	low_avg_freight_value_state	low_average_freight
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

Job history REFRESH

<u>Insights</u>:

1. The states with the highest average freight values like states called RR and PB may experience greater shipping prices due to reasons like remote locations, higher transportation costs, or supply chain difficulties.

2. It might be useful for our company to try to optimize logistics operations or save costs to locate places with relatively reduced shipping prices by looking at the states with the lowest average freight values like states such as SP and PR.

<u>Recommendations</u>:

1.Implement strategic cost management initiatives for states with higher average freight values, such as RR and PB. This could involve negotiating freight costs, optimizing logistics operations, or exploring alternative transportation solutions to mitigate the impact of remote locations or higher transportation costs.
2.Develop focused cost reduction initiatives based on the insights gained from the data. This might include negotiating favorable terms with carriers, exploring bulk shipping discounts, or identifying supply chain efficiencies to minimize shipping costs.

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

Query:

```
with cte as (
select
        c.customer_state as state,
        avg(t.delivery_time) as avg_delivery
from (
SELECT *,
        ROUND(TIMESTAMP_DIFF(order_delivered_customer_date,
        order_purchase_timestamp, day),2) AS delivery_time,
FROM
        `Project1TargetSQL.orders`
```

```
WHERE
        order_status = 'delivered' AND order_delivered_customer_date IS NOT NULL
ORDER BY
        order_purchase_timestamp
) t
JOIN
         `Project1TargetSQL.customers`  c
ON
        t.customer_id = c.customer_id

group by
        c.customer_state
),

high_state as (
select
         state as high_state,
        avg_delivery as high_avg_delivery,
         row_number() over (order by avg_delivery desc ) as rnk1
from cte ),

low_state as (
        select
                state as low_state,
                avg_delivery as low_avg_delivery,
                row_number() over (order by avg_delivery asc ) as rnk2
from cte )
select
         high_state,
         round(high_avg_delivery,2) as high_avg_delivery,
         low_state,
```

```
        round(low_avg_delivery,2) as low_avg_delivery
from
        high_state hs
join
        low_state ls
on
        hs.rnk1 = ls.rnk2
limit 5
```

Output screenshot:

Query results

SAVE RESULTS EXPLORE DATA

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	high_state	high_avg_delivery	low_state	low_avg_delivery
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

Insights:

1. Finding areas with effective delivery operations, quicker transit times, or solid logistics networks can be done by looking at the states like SP and PR with the lowest average delivery times and states called RR and AP with highest average delivery times.

2. These insights can be helpful for our company looking to improve customer satisfaction, operational efficiency, delivery process optimization, and setting reasonable expectations for customers based on regional delivery time patterns.

Recommendations:

1.Collaborate with local stakeholders, including carriers and distribution partners, to optimize delivery operations in specific regions. Local insights and partnerships can contribute to more effective and tailored solutions.
2.Utilize insights to improve customer satisfaction by setting realistic delivery expectations based on regional patterns. Communicate transparently with customers about delivery times, and consider implementing initiatives to further enhance the delivery process.

4. 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:

```
with cte as (
select
        customer_state as state,
        avg(date_diff(order_delivered_customer_date,
        order_estimated_delivery_date,day)) as avg_delivery_speed,
        row_number() over (order by avg(
        date_diff(order_delivered_customer_date,order_estimated_delivery_date,
        day)) asc) as rnk
from
        `Project1TargetSQL.orders` o
join
        `Project1TargetSQL.customers` c
on
        o.customer_id = c.customer_id
WHERE
        o.order_delivered_customer_date IS NOT NULL AND
        o.order_estimated_delivery_date IS NOT NULL
group by
        customer_state
)
select
        state ,
        avg_delivery_speed
from cte
where rnk <= 5
```

Output screenshot:

Query results

SAVE RESULTS EXPLORE DATA

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	state	avg_delivery_speed
1	AC	-19.7625000000...
2	RO	-19.1316872427...
3	AP	-18.7313432835...
4	AM	-18.6068965517...
5	RR	-16.4146341463...

Insights:

1. Company operating in these states called AC, RO, AP, and AM where average delivery speed is highest as negative indicates it deliver before the actual date of delivery. We can take advantage of the quicker delivery times by highlighting their rapid and dependable service, thereby drawing more clients, and boosting client satisfaction.

2. These data can help us improve our operations, enhance customer experience, optimize logistics, or look for expansion prospects in areas with a track record of quick order delivery.

Recommendations:

1. Highlighting the rapid and dependable service in marketing campaigns can attract more clients and significantly boost customer satisfaction.
2.Position the brand as synonymous with speed and reliability. Incorporate the quick delivery times into the company's brand messaging, emphasizing the commitment to efficient service and meeting or exceeding customer expectations.

6. **Analysis based on the payments:**
   1. Find the month on month no. of orders placed using different payment types.

Query:

```
select
        format_timestamp('%Y-%m',o.order_purchase_timestamp) as Months,
        p.payment_type ,
        count(distinct o.order_id) as Number_of_orders
from
        `Project1TargetSQL.orders` o
join
        `Project1TargetSQL.payments` p
on
        o.order_id =p.order_id
group by
```

```
        Months,
        p.payment_type
order by
        Months
```

Output screenshot:

Query results

SAVE RESULTS EXPLORE DATA

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	Months	payment_type	Number_of_orders
1	2016-09	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	2017-01	credit_card	582
8	2017-01	UPI	197
9	2017-01	voucher	33
10	2017-01	debit_card	9

Results per page: 50 1 – 50 of 90

Insights:

1.We identify credit card as a payment method was mostly used followed by UPI , Vouchers , debit card then other.

2. Credit card as a payment method was most used in November 2017

3.Based on the payment preferences noticed during various months, these insights might help firms optimize their payment procedures, customize marketing campaigns, or enhance customer experiences.

Recommendations:

1.Customize marketing campaigns based on the popularity of payment methods. Tailor promotions, discounts, or loyalty programs that specifically target users of credit cards, UPI, Vouchers, and debit cards to maximize engagement and conversion.
2.Explore opportunities for promotional offers and partnerships with financial institutions or payment service providers. Collaborate to offer exclusive discounts, cashback, or incentives for users of popular payment methods, thereby encouraging customer loyalty and increasing transaction volumes.

2. 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 order_count
from
        `Project1TargetSQL.payments`
where
        payment_installments > 0
group by
        payment_installments
order by
        payment_installments
```

Output screenshot:

Query results

SAVE RESULTS EXPLORE DATA

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	payment_installment	order_count
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

Results per page: 50 1 – 23 of 23

Insights:

1.The majority of orders (49,060) are made with a single instalment, indicating that a significant portion of customers prefers or has the financial capacity to make one-time payments for their orders.

2.As the number of instalments increases, the order count gradually decreases. This suggests that a decreasing proportion of customers opt for multiple instalments, with fewer customers choosing to spread payments across several periods.

Recommendations:

1.Launch targeted promotions to encourage customers to opt for multiple instalments. Consider offering discounts, cashback, or exclusive incentives for orders with two or more instalments to attract a larger audience.
2.Offer flexible instalment plans that cater to varying financial capacities. Provide options for 2, 3, or 4 instalments to accommodate different budget preferences,

ensuring that customers can choose plans that align with their financial circumstances.