TARGET DATA

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

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

```sql
SELECT
 table_name,
 column_name,
 data_type
FROM Target.INFORMATION_SCHEMA.COLUMNS
WHERE table_name = 'customers'
```

Query results

	JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH PR

Row	table_name ▼	column_name ▼	data_type ▼
1	customers	customer_id	STRING
2	customers	customer_unique_id	STRING
3	customers	customer_zip_code_prefix	INT64
4	customers	customer_city	STRING
5	customers	customer_state	STRING

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

```sql
SELECT
 TIMESTAMP_DIFF(max(order_purchase_timestamp),min(order_purchase_timestamp),DAY) as
Date_difference,
 TIMESTAMP_DIFF(max(order_purchase_timestamp),min(order_purchase_timestamp),HOUR)
as Time_difference,
FROM `Target.orders`
```

Query results

	JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS

Row	Date_difference ▼	Time_difference ▼	
1	772	18548	

C. Count the number of Cities and States in our dataset.

```sql
SELECT
 COUNT(DISTINCT geolocation_city) AS city_count,
 COUNT(DISTINCT geolocation_state) AS state_count
FROM `Target.geolocation`
```

Query results

	JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS

Row	city_count ▼	state_count ▼	
1	8011	27	

INSIGHTS:-

The presence of customers across 27 unique states and 8,011 unique cities indicates that Target has established a wide geographic reach in the Brazil region. This suggests a strong market presence and a broad customer base spread throughout the country.

In-depth Exploration:

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

```sql
SELECT
    EXTRACT(YEAR from order_purchase_timestamp) as Year,
    EXTRACT(MONTH from order_purchase_timestamp) as Month,
    COUNT(order_id)as No_of_orders
  FROM `Target.orders`
  GROUP BY Year,Month
  ORDER BY Year,Month;
```

Query results

	JOB INFORMATION		RESULTS		JSON		EXECUTION DETAILS	

Row	Year ▾	Month ▾	No_of_orders ▾
1	2016	9	4
2	2016	10	324
3	2016	12	1
4	2017	1	800
5	2017	2	1780
6	2017	3	2682
7	2017	4	2404
8	2017	5	3700
9	2017	6	3245
10	2017	7	4026

INSIGHTS:-

Overall Order Growth: This analysis reveals a consistent and significant upward trend in the number of orders over the past years in the Brazil region. This indicates a positive business performance and growing customer demand for the Target company's products in Brazil.

No_of_orders and Month
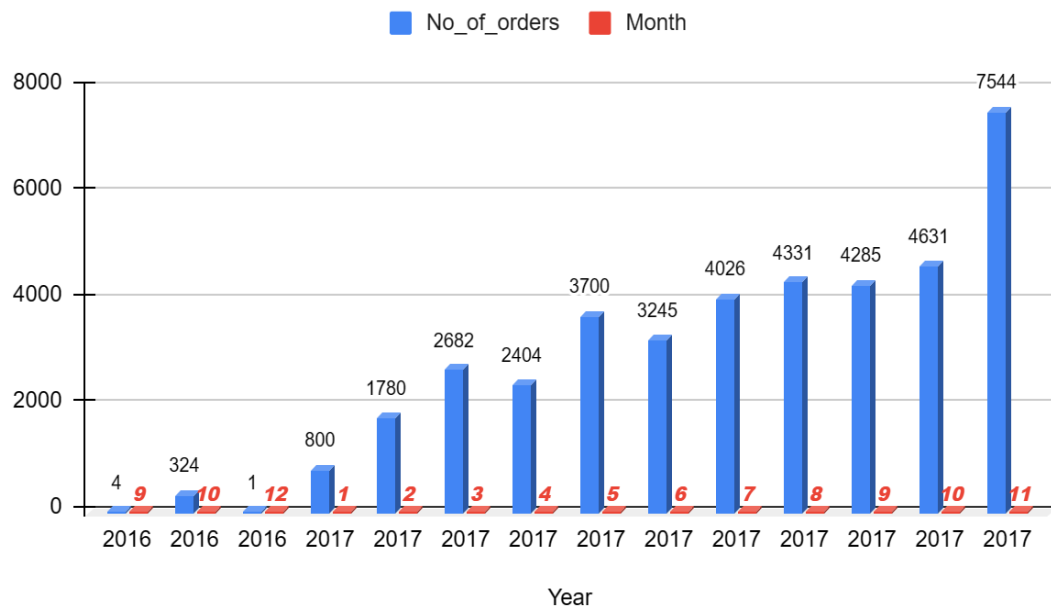
B. Can we see some kind of monthly seasonality in terms of the no. of orders being placed?

```sql
SELECT
    EXTRACT(MONTH from order_purchase_timestamp) as Month,
    COUNT(order_id)as No_of_orders
  FROM `Target.orders`
  GROUP BY Month
  ORDER BY Month;
```

JOB INFORMATION	RESULTS	JSON

Row	Month ▼	No_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:-

Seasonal trends: It appears that there may be some seasonal trends in the orders. The number of orders seems to increase from January (8069) to March (9893), and then decrease from March to April (9343). There is a subsequent increase from April to May (10573) and a slight decrease from May to June (9412). The number of orders is relatively lower in September (4305) and October (4959), compared to other months. The lowest number of orders is in September (4305), and the highest is in August (10843).
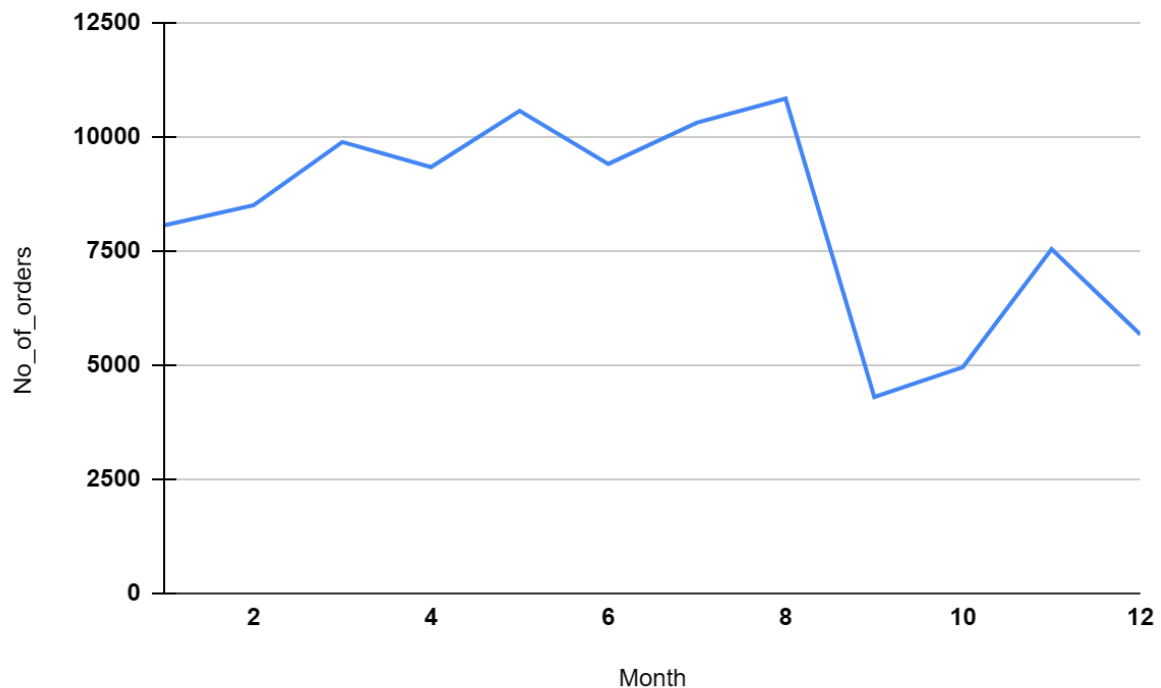
RECOMMENDATIONS:-

- ➢ Conduct market research to understand the reasons behind the low-order months. Analyze customer behavior, competitor activity, and market trends during those periods.

- ➢ Explore collaborations or partnerships with complementary businesses. By cross-promoting each other's products or services, you can tap into a broader customer base and potentially increase orders during the slower

- ➢ Offer special promotions, discounts, or limited-time offers during the low-order months to incentivize customers to make purchases. This could include seasonal discounts, buy-one-get-one (BOGO) offers, or free shipping for a limited period.

- ➢ Now it's a world of E-commerce, so Consider utilizing email marketing, social media promotions, and paid advertising to reach a wider audience on those particular months.

C. 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**

```sql
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 Time_of_day,
  COUNT(*) AS No_of_orders
FROM
  `Target.orders`
GROUP BY
  time_of_day
ORDER BY
  No_of_orders DESC;
```

Query results

JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS

Row	Time_of_day ▼	No_of_orders ▼
1	Afternoon	38135
2	Night	28331
3	Morning	27733
4	Dawn	5242

INSIGHTS:-

Afternoon (38135 orders): The highest number of orders seems to occur during the afternoon. This could be attributed to several factors such as people being more active and engaged during the daytime, taking breaks from work or other activities, and having more time for browsing and making purchases.

Night (28331 orders): The number of orders during the night is slightly lower compared to the afternoon. It is common for online shopping to continue into the evening hours when people are at home and have more leisure time. However, the decrease in order volume could be due to factors such as fatigue, reduced browsing time, or engaging in other nighttime activities.

Morning (27733 orders): The number of orders during the morning is relatively close to the nighttime orders. This suggests that some customers prefer to make their purchases early in the day, potentially before starting their work or daily activities. However, it's important to note that morning orders might vary depending on the target audience and industry.

Dawn (5242 orders): The lowest number of orders occurs during the dawn hours. This could be due to the fact that dawn is typically the early morning period when people are still asleep or just starting their day. During this time, there may be fewer people engaged in online shopping activities.

RECOMMENDATIONS:-

Afternoon :

Run time-limited promotions, offer personalized marketing, and ensure responsive customer support to capitalize on the high order volume in the afternoon.

Night :

Optimize for mobile, engage on social media, and create nighttime promotions to capture the attention of users during the evening hours.

Morning :

Offer exclusive early bird discounts, leverage email marketing, and showcase social proof to incentivize morning shoppers to make purchases.

Dawn :

Implement automated marketing campaigns, create time-sensitive offers, and gather customer feedback to leverage the dawn hours for increased conversions.

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

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

```sql
SELECT
    c.customer_state,
    EXTRACT(MONTH FROM order_purchase_timestamp) AS Month,
    COUNT(order_id) AS No_of_orders
FROM `Target.orders` AS o
LEFT JOIN `Target.customers` AS c
ON o.customer_id = c.customer_id
GROUP BY Month, c.customer_state
ORDER BY c.customer_state, Month;
```

JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS

Row	customer_state ▾	Month ▾	No_of_orders ▾
1	AC	1	8
2	AC	2	6
3	AC	3	4
4	AC	4	9
5	AC	5	10
6	AC	6	7
7	AC	7	9
8	AC	8	7
9	AC	9	5
10	AC	10	6

INSIGHTS:-

By examining the month-on-month order data, we can identify trends and patterns that can aid in forecasting future order volumes. This information can support production planning, inventory management, and resource allocation to meet anticipated demand in different states.

For example:- Take the "AC" state. The number of orders in state "AC" varies throughout the year, with different levels of activity in different months. The highest number of orders is observed in Month 5 (10 orders), followed by Months 4 and 7 with 9 orders each. On the other hand, the lowest number of orders is observed in Months 9, 11, and 12, with 5 orders each. here is no clear trend in order volumes across the months for state "AC".

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

```sql
SELECT
    customer_state,
    COUNT(customer_unique_id) AS No_of_unique_customers
FROM `Target.customers`
GROUP BY customer_state
ORDER BY customer_state;
```

JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS

Row	customer_state ▾	No_of_unique_customers
1	AC	81
2	AL	413
3	AM	148
4	AP	68
5	BA	3380
6	CE	1336
7	DF	2140
8	ES	2033
9	GO	2020
10	MA	747

INSIGHTS:-

 The data indicates that different states have varying numbers of unique customers. States like São Paulo (SP) and Rio de Janeiro (RJ) have a significantly higher number of unique customers, with SP having the highest count of 41,746 and RJ having 12,852. On the other hand, states like Roraima (RR) and Tocantins (TO) have a relatively lower number of unique customers, with RR having the lowest count of 46 and TO having 280.

RECOMMENDATIONS:-

> Focus on Customer Engagement: In states with higher customer bases like São Paulo (SP) and Rio de Janeiro (RJ), prioritize customer engagement and retention strategies to strengthen relationships and encourage repeat purchases.

> Target Growth Opportunities: States with lower customer bases such as Roraima (RR) and Tocantins (TO) represent growth potential. Expand marketing efforts, tailor messages to local preferences, and offer incentives to attract new customers.

> Identify and leverage your competitive advantage in states with higher customer bases. Analyze competitors' strategies and positioning to differentiate your brand and provide unique value propositions. Highlight factors such as quality, pricing, convenience, or exclusive product offerings to stand out in the market.

IV. 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).

```
SELECT ROUND(((total_cost_2018 - total_cost_2017) / total_cost_2017) * 100) AS
percentage_increase
FROM
 (SELECT
   (SELECT SUM(payment_value)
    FROM `Target.payments` as p
    JOIN `Target.orders` as o ON p.order_id = o.order_id
    WHERE EXTRACT(YEAR FROM o.order_purchase_timestamp) = 2017
      AND EXTRACT(MONTH FROM o.order_purchase_timestamp) BETWEEN 1 AND 8) AS
total_cost_2017,
   (SELECT SUM(payment_value)
    FROM `Target.payments` as p
    JOIN `Target.orders` as o ON p.order_id = o.order_id
    WHERE EXTRACT(YEAR FROM o.order_purchase_timestamp) = 2018
      AND EXTRACT(MONTH FROM o.order_purchase_timestamp) BETWEEN 1 AND 8) AS
total_cost_2018 ) AS costs
```

Query results

JOB INFORMATION	RESULTS	JSON

Row	percentage_increase
1	137.0

INSIGHTS:-

The substantial increase in order costs indicates a significant pricing adjustment or changes in product offerings during that period. External factors like inflation, raw material costs, and demand-supply dynamics likely influenced the cost increase. Understanding these market dynamics is crucial to contextualize the change.

RECOMMENDATIONS:-

➢ Analyze customer response to the cost increase. Higher prices could have affected customer loyalty and purchasing decisions, warranting attention to customer feedback and market research.

➢ Compare the cost increase with industry competitors to identify if it aligns with market trends or diverges significantly.

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

```sql
SELECT
 customer_state,
 ROUND(sum(o.price),0) AS total_price,
 ROUND(avg(o.price),0) AS average_price
FROM `Target.order_items`AS o
JOIN `Target.orders` AS od
ON o.order_id = od.order_id
JOIN `Target.customers` AS c
ON od.customer_id = c.customer_id
GROUP BY
 customer_state
ORDER BY customer_state;
```

JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS

Row	customer_state	total_price	average_price
1	AC	15983.0	174.0
2	AL	80315.0	181.0
3	AM	22357.0	135.0
4	AP	13474.0	164.0
5	BA	511350.0	135.0
6	CE	227255.0	154.0
7	DF	302604.0	126.0
8	ES	275037.0	122.0
9	GO	294592.0	126.0
10	MA	119648.0	145.0

INSIGHTS:-

There is significant variation in total and average order prices across different states. For example, states like São Paulo (SP) and Rio de Janeiro (RJ) have higher total prices, indicating larger revenue contributions, while states like Roraima (RR) and Amapá (AP) have relatively lower total prices.

Average order prices reflect the purchasing power or willingness of customers in each state. States like Amapá (AP) and Roraima (RR) have higher average order prices, suggesting potential higher-value customers or demand for premium-priced products. On the other hand, states like Espírito Santo (ES) and Goiás (GO) have relatively lower average order prices.

RECOMMENDATIONS:-

➢ States with higher total and average order prices, such as São Paulo (SP), Rio de Janeiro (RJ), and Minas Gerais (MG), represent high-value markets. Allocate additional resources and marketing efforts to these states to capitalize on their revenue potential.

➢ Identify states with lower total and average order prices, such as Espirito Santo (ES), Goias (GO), and Para (PA), and conduct further analysis to understand the reasons behind the relatively lower order values. Evaluate market dynamics, customer behaviour, competition, and pricing strategies in these states. Consider adjusting pricing, improving product offerings, or implementing targeted marketing campaigns to stimulate demand and increase order values.

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

```
SELECT
 customer_state,
 ROUND(sum(o.freight_value),0) AS total_freight_value,
 ROUND(avg(o.freight_value),0) AS average_freight_value
FROM `Target.order_items`AS o
JOIN `Target.orders` AS od
ON o.order_id = od.order_id
JOIN `Target.customers` AS c
ON od.customer_id = c.customer_id
GROUP BY
 customer_state
ORDER BY customer_state;
```

JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS

Row	customer_state ▾	total_freight_value	average_freight_valu
1	AC	3687.0	40.0
2	AL	15915.0	36.0
3	AM	5479.0	33.0
4	AP	2789.0	34.0
5	BA	100157.0	26.0
6	CE	48352.0	33.0
7	DF	50625.0	21.0
8	ES	49765.0	22.0
9	GO	53115.0	23.0
10	MA	31524.0	38.0

INSIGHTS:-

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 in each state. States like São Paulo (SP) and Minas Gerais (MG) have lower average freight values, suggesting potentially more efficient logistics networks or economies of scale due to higher order volumes. In contrast, states like Paraíba (PB) and Piauí (PI) have higher average freight values, indicating relatively higher shipping costs per order

RECOMMENDATIONS:-

➢ Analyze states with higher average freight values, such as Paraíba (PB) and Piauí (PI), to identify potential areas for optimizing shipping strategies. Consider negotiating better shipping rates with carriers, exploring alternative logistics providers, or implementing order consolidation techniques to reduce shipping costs

➢ Assess states with higher total freight values, such as São Paulo (SP) and Minas Gerais (MG), to identify opportunities for improving operational efficiency. Streamline logistics processes, such as warehouse management, order fulfillment, and route planning, to minimize shipping costs. Explore partnerships with local distribution centers or fulfillment services to optimize order fulfillment and reduce transit distances. By enhancing operational efficiency, you can lower shipping costs and improve overall logistics performance.

V. 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.

```sql
SELECT
  order_id,
  EXTRACT(DATE FROM order_purchase_timestamp) AS order_purchase_date,
  EXTRACT(DATE FROM order_delivered_customer_date) AS order_delivered_date,
  EXTRACT(DATE FROM order_estimated_delivery_date) AS estimated_delivery_date,
  DATE_DIFF(EXTRACT(DATE FROM order_delivered_customer_date), EXTRACT(DATE FROM order_purchase_timestamp), DAY) AS time_to_deliver,
  DATE_DIFF(EXTRACT(DATE FROM order_estimated_delivery_date), EXTRACT(DATE FROM order_delivered_customer_date), DAY) AS diff_estimated_delivery
FROM `Target.orders`
WHERE
  order_purchase_timestamp IS NOT NULL
  AND order_delivered_customer_date IS NOT NULL
  AND order_estimated_delivery_date IS NOT NULL
ORDER BY order_purchase_date;
```

Row	order_id ▾	order_purchase_date	order_delivered_date	estimated_delivery_c	time_to_deliver ▾	diff_estimated_delivery
1	bfbd0f9bdef84302105ad712db...	2016-09-15	2016-11-09	2016-10-04	55	-36
2	65d1e226dfaeb8cdc42f66542...	2016-10-03	2016-11-08	2016-11-25	36	17
3	be5bc2f0da14d8071e2d45451...	2016-10-03	2016-10-27	2016-11-07	24	11
4	ae8a60e4b03c5a4ba9ca0672c...	2016-10-03	2016-11-03	2016-12-01	31	28
5	cd3b8574c82b42fc8129f6d50...	2016-10-03	2016-10-14	2016-11-23	11	40
6	d207cc272675637bfed0062ed...	2016-10-03	2016-10-31	2016-11-23	28	23
7	a41c8759fbe7aab36ea07e038...	2016-10-03	2016-11-03	2016-11-29	31	26
8	ef1b29b591d31d57c0d733746...	2016-10-03	2016-11-01	2016-11-25	29	24
9	3b697a20d9e427646d925679...	2016-10-03	2016-10-26	2016-10-27	23	1
10	35d3a51724a47ef1d0b89911e...	2016-10-04	2016-10-26	2016-12-20	22	55

INSIGHTS:-

From the above result the "Time_to_deliver" column represents the no.of.days taken to deliver an order to the customer from thr oreder purchased date. The "diff_estimated_delivery" is the difference between the estimated delivery date and order delivered date to the customer.

We can find some negative values in this coloum it represents that the order is delivered after the estimated delivery date, if the value is in plus it indicates the order is delivered prior the estimated delivery date

For example:

Take the first order it has minus 36 days in the diff_estimated_delivery coloumn it shows that the order has been deliverd 36 days after the estimsted delivery date.

Take the second order it has plus 17 days in the diff_estimated_delivery coloumn it shows that the order is delivered in 17 days prior to the estimated delivery date.

RECOMMENDATIONS:-

➢ Aim to minimize the time-to-delivery by optimizing logistics processes, improving inventory management, and streamlining order fulfillment operations.

➢ Evaluate the accuracy of estimated delivery times to ensure they align with the actual delivery performance. Identify any discrepancies and work towards improving the estimation process to provide customers with more reliable delivery estimates.

➢ Review the diff_estimated_delivery and estimated_delivery_date columns to identify cases where orders were delivered later than the estimated delivery date. Investigate the reasons behind these delays and identify any bottlenecks or operational issues that may be causing delays in order fulfillment.

➢ By efficiently managing resources and capacity, you can ensure timely order fulfillment and maintain high customer service levels. This can be done by analyzing the order_purchase_date column

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

```sql
WITH avg_freight AS (
  SELECT
    c.customer_state,
    ROUND(AVG(o.freight_value), 0) AS average_freight_value,
    ROW_NUMBER() OVER (ORDER BY AVG(o.freight_value)) AS rn_asc,
    ROW_NUMBER() OVER (ORDER BY AVG(o.freight_value) DESC) AS rn_desc
  FROM `Target.order_items` AS o
  JOIN `Target.orders` AS od ON o.order_id = od.order_id
  JOIN `Target.customers` AS c ON od.customer_id = c.customer_id
  GROUP BY
    c.customer_state
)
SELECT
  CASE
    WHEN rn_desc <= 5 THEN 'Top 5 states with Highest_avg_freight_value'
    WHEN rn_asc <= 5 THEN 'Top 5 states with lowest_avg_freight_value'
    ELSE ''
  END AS freight_category,
   customer_state,
  average_freight_value
FROM avg_freight
WHERE rn_asc <= 5 OR rn_desc <= 5
ORDER BY average_freight_value;
```

| JOB INFORMATION | RESULTS | JSON | EXECUTION DETAILS | EXECUTION G |

Row	freight_category ▼	customer_state ▼	average_freight_valu
1	Top 5 states with lowest_avg_f...	SP	15.0
2	Top 5 states with lowest_avg_f...	DF	21.0
3	Top 5 states with lowest_avg_f...	RJ	21.0
4	Top 5 states with lowest_avg_f...	MG	21.0
5	Top 5 states with lowest_avg_f...	PR	21.0
6	Top 5 states with Highest_avg_...	PI	39.0
7	Top 5 states with Highest_avg_...	AC	40.0
8	Top 5 states with Highest_avg_...	RO	41.0
9	Top 5 states with Highest_avg_...	RR	43.0
10	Top 5 states with Highest_avg_...	PB	43.0

RECOMMENDATIONS:-

➢ 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), consider implementing strategies to further optimize freight costs. Explore opportunities to negotiate better shipping rates with carriers, consolidate shipments to reduce costs, and optimize logistics routes to minimize distances and improve efficiency. By focusing on cost-saving measures, you can maintain competitive pricing and potentially attract more customers in these states.

➢ Highest Average Freight Value States: In states with the highest average freight values, such as Piauí (PI), Acre (AC), Rondônia (RO), Roraima (RR), and Paraíba (PB), it is essential to evaluate and optimize the freight value structure. Identify the factors contributing to higher freight costs, such as longer distances, remote locations, or specific product characteristics, and seek opportunities to mitigate these factors. Explore partnerships with local logistics providers, optimize transportation routes, and leverage technology for efficient freight management.

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

```sql
WITH avg_freight AS (
  SELECT
    c.customer_state,
    ROUND(AVG(DATE_DIFF((extract(DATE FROM
order_delivered_customer_date)),(extract(DATE FROM
order_purchase_timestamp)),Day)),0)as Avg_Delivery_time,
    ROW_NUMBER() OVER (ORDER BY AVG(DATE_DIFF((extract(DATE FROM
order_delivered_customer_date)),(extract(DATE FROM
order_purchase_timestamp)),Day))) AS rn_asc,
    ROW_NUMBER() OVER (ORDER BY AVG(DATE_DIFF((extract(DATE FROM
order_delivered_customer_date)),(extract(DATE FROM order_purchase_timestamp)),Day))
DESC) AS rn_desc
  FROM `Target.order_items` AS o
  JOIN `Target.orders` AS od ON o.order_id = od.order_id
  JOIN `Target.customers` AS c ON od.customer_id = c.customer_id
  GROUP BY
    c.customer_state
)
SELECT
  CASE
    WHEN rn_desc <= 5 THEN 'Top 5 states with Highest_Avg_Delivery_time'
    WHEN rn_asc <= 5 THEN 'Top 5 states with lowest_Avg_Delivery_time'
    ELSE ''
  END AS Delivery_category,
```

```
    customer_state,
    Avg_Delivery_time
FROM avg_freight
WHERE rn_asc <= 5 OR rn_desc <= 5
ORDER BY Avg_Delivery_time;
```

	JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH

Row	Delivery_category ▼	customer_state ▼	Avg_Delivery_time
1	Top 5 states with lowest_Avg_Delivery_time	SP	9.0
2	Top 5 states with lowest_Avg_Delivery_time	MG	12.0
3	Top 5 states with lowest_Avg_Delivery_time	PR	12.0
4	Top 5 states with lowest_Avg_Delivery_time	DF	13.0
5	Top 5 states with lowest_Avg_Delivery_time	SC	15.0
6	Top 5 states with Highest_Avg_Delivery_time	AL	24.0
7	Top 5 states with Highest_Avg_Delivery_time	PA	24.0
8	Top 5 states with Highest_Avg_Delivery_time	AM	26.0
9	Top 5 states with Highest_Avg_Delivery_time	AP	28.0
10	Top 5 states with Highest_Avg_Delivery_time	RR	28.0

RECOMMENDATIONS:-

➢ In states with the highest average delivery times, such as Alagoas (AL), Pará (PA), Amazonas (AM), Amapá (AP), and Roraima (RR), it is crucial to identify and address the factors contributing to these extended delivery times. Evaluate your logistics network, transportation routes, and last-mile delivery processes to optimize efficiency. Explore opportunities to collaborate with local logistics providers or establish strategic partnerships to improve delivery performance in these states. By reducing delivery times, you can enhance customer experience and mitigate any potential negative impact on customer satisfaction.

➢ Analyze the last-mile delivery process in each state to identify potential bottlenecks or areas for improvement. Consider implementing technology solutions such as route optimization software, real-time tracking systems, and efficient delivery scheduling tools to streamline the last-mile operations. By optimizing the last-mile delivery process, you can enhance delivery efficiency, reduce transit times, and improve the overall customer experience.

➢ Maintain open communication and collaborative relationships with shipping carriers and logistics partners. Regularly review performance metrics, including delivery times, and address any concerns or issues promptly. Work together to identify areas for improvement, streamline processes, and optimize transportation routes to ensure faster and more reliable deliveries across all states

D. Find out the top 5 states where the order delivery is really fast as compared to the estimated date of delivery.

```sql
WITH avg_delivery AS (
  SELECT
    c.customer_state,
    ROUND((AVG(DATE_DIFF(DATE(order_estimated_delivery_date),
DATE(order_delivered_customer_date), DAY))),0) as Avg_diff_estimated_delivery,
    ROW_NUMBER() OVER (ORDER BY AVG(DATE_DIFF(DATE(order_estimated_delivery_date),
DATE(order_delivered_customer_date), DAY))DESC) AS top_fast_delivery
  FROM `Target.order_items` AS o
  JOIN `Target.orders` AS od ON o.order_id = od.order_id
  JOIN `Target.customers` AS c ON od.customer_id = c.customer_id
  WHERE order_delivered_customer_date IS NOT NULL
  GROUP BY
    c.customer_state )
SELECT
  customer_state as Top_5_fast_delivery_states,
  Avg_diff_estimated_delivery
FROM avg_delivery
WHERE top_fast_delivery <= 5
ORDER BY Avg_diff_estimated_delivery DESC;
```

JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS

Row	Top_5_fast_delivery_states ▼	Avg_diff_estimated_delivery ▼
1	AC	21.0
2	RO	20.0
3	AM	20.0
4	AP	18.0
5	RR	18.0

INSIGHTS:-

These states have an average difference of 18 to 21 days between the estimated delivery date and the actual delivery date. There may be well-established transportation networks, optimized routing strategies, and proactive coordination with shipping carriers.

RECOMMENDATIONS:-

 Maintain clear and proactive communication with customers regarding delivery expectations. Provide accurate and timely updates on order statuses, shipment tracking, and any potential delays to manage customer expectations effectively.

VI. Analysis based on the payments:

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

```sql
SELECT
EXTRACT(MONTH FROM order_purchase_timestamp) as Month,
payment_type,
count(p.order_id) as No_of_orders
FROM `Target.payments` as p
JOIN `Target.orders` as o ON
p.order_id = o.order_id
GROUP BY Month,payment_type
ORDER BY Month,payment_type;
```

	JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	E)
Row	Month ▼		payment_type ▼		No_of_orders ▼	
1		1	UPI		1715	
2		1	credit_card		6103	
3		1	debit_card		118	
4		1	voucher		477	
5		2	UPI		1723	
6		2	credit_card		6609	
7		2	debit_card		82	
8		2	voucher		424	
9		3	UPI		1942	
10		3	credit_card		7707	

INSIGHTS:-

Credit card emerges as the preferred payment method, possibly due to its convenience, security, and widespread acceptance. It is crucial to ensure seamless credit card payment processing and maintain strong partnerships with payment service providers to facilitate smooth transactions.

RECOMMENDATIONS:-

➢ Ensure a robust and secure payment infrastructure that supports seamless transactions across various payment methods. Regularly update payment systems to stay aligned with evolving customer preferences and emerging payment technologies.

➢ Educate customers about the benefits and convenience of using alternative payment methods such as UPI, debit cards, and vouchers. Offer incentives or discounts to encourage customers to explore and use these options.

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

```sql
SELECT
payment_installments,
count(order_id) as No_of_Orders
FROM `Target.payments`
WHERE payment_installments <> 0
GROUP BY payment_installments
```

JOB INFORMATION	RESULTS	JSON	EXE

Row	payment_installments	No_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

INSIGHTS:-

The data reflects a significant customer preference for dividing payments into multiple installments, particularly around the 9-10 installment range. Businesses should consider offering flexible payment options and installment plans to cater to customer preferences and increase sales.

RECOMMENDATIONS:-

➢ Run targeted promotional campaigns to highlight the benefits of choosing lower installment options, such as affordability, reduced financial burden, and ease of managing payments. Leverage marketing channels to educate customers about the advantages of installment plans and create a sense of value in choosing lower installments.

➢ Regularly gather customer feedback and analyze customer behavior to gain insights into their preferences and pain points related to installment payments.