Business Case: Target SQL

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 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.

#Q1 Import the dataset and do usual exploratory analysis steps like checking the structure & characteristics of the dataset:

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

Answer Query:

```
SELECT COLUMN_NAME, DATA_TYPE
FROM `artful-striker-417618.target_sql.INFORMATION_SCHEMA.COLUMNS`
WHERE TABLE_NAME = 'customers';
```

Output Screenshot:

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:

The dataset contains customer data such as unique identifiers, zip code prefixes, city names, and state information. The string data types indicate textual data, while the integer data type indicates numerical identifiers.

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

Answer Query:

```
SELECT  min(order_purchase_timestamp) as first_order_date,
        max(order_purchase_timestamp) as last_order_date
FROM  `target_sql.orders`;
```

Output Screenshot:

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

Insights:

The dataset's order placements span 2 years and 1 month, starting on September 4, 2016, and ending on October 17, 2018, providing valuable context for understanding the temporal distribution of orders.

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

Answer Query:

```
SELECT count(DISTINCT c.customer_city) AS number_of_cities,
       count(DISTINCT c.customer_state) AS number_of_states
FROM `target_sql.customers` AS c
INNER JOIN `target_sql.orders` AS o ON c.customer_id = o.customer_id;
```

Output Screenshot:

Row	number_of_cities	number_of_states
1	4119	27

Insights:

During the specified period, 4,119 unique cities and 27 states were the locations where customers placed orders, indicating a wide geographical reach and a broad customer base across various regions, contributing to a rich and varied dataset.

#Q2 In-depth Exploration:

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

Answer Query:

```
SELECT  EXTRACT(YEAR FROM order_purchase_timestamp) AS year_of_purchase,
        COUNT(*) AS number_of_orders
FROM   `target_sql.orders`
GROUP BY  EXTRACT(YEAR FROM order_purchase_timestamp)
ORDER BY  year_of_purchase;
```

Output Screenshot:

Row	year_of_purchase	number_of_orders
1	2016	329
2	2017	45101
3	2018	54011

Insights:

The number of orders has been increasing year over year.

- In 2016, there were 329 orders.
- The orders significantly jumped to 45,101 in 2017.
- The growth continued in 2018 with 54,011 orders.

This indicates a positive and growing trend in the number of orders placed over the years. The business seems to be expanding, attracting more customers and orders each year. This is a good sign of growth and customer engagement.

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

Answer Query:

```
SELECT EXTRACT(YEAR FROM order_purchase_timestamp) AS year_of_purchase,
       EXTRACT(MONTH FROM order_purchase_timestamp) AS month_of_purchase,
       COUNT(*) AS number_of_orders
FROM
  `target_sql.orders`
GROUP BY
  1, 2
ORDER BY
  1,2;
```

Output Screenshot:

Row	year_of_purchase	month_of_purchase	number_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:

In 2017, the number of orders increased steadily from August to November, with a peak in November. This could be attributed to the holiday shopping season, including Black Friday and Cyber Monday events. However, there was a slight decrease in December due to the end of the holiday season.

In 2018, the number of orders remained relatively high from January to August, indicating a strong performance throughout most of the year. However, there was a significant drop in September and October, possibly due to the FIFA World Cup 2018, which typically features promotional activities and special offers.

The data reveals a clear seasonal pattern, with increases during certain periods like the holiday season and the FIFA World Cup. It's important for businesses to understand these seasonal trends to plan their marketing and sales strategies effectively.

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

Answer 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 'Mornings'
  WHEN EXTRACT(HOUR FROM order_purchase_timestamp) BETWEEN 13 AND 18 THEN 'Afternoon'
  ELSE 'Night' end as Time_Of_Day,
  count(*)as Orders_Placed_Count
FROM
  `target_sql.orders`
GROUP BY
  1
ORDER BY
  2 DESC;
```

Output Screenshot:

Row	Time_Of_Day	Orders_Placed_Coun
1	Afternoon	38135
2	Night	28331
3	Mornings	27733
4	Dawn	5242

Insights: The data suggests that the majority of orders are placed in the afternoon, followed by night and mornings. Very few orders are placed at dawn. This could be due to customer habits or the nature of the products being ordered.

Understanding these patterns can help in planning marketing campaigns, optimizing operations, and improving customer service.

#Q3 Evolution of E-commerce orders in the Brazil region:

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

Answer Query:

```
SELECT c.customer_state,
       FORMAT_TIMESTAMP('%Y-%m', o.order_purchase_timestamp) AS Year_Month,
       COUNT(*) AS number_of_orders
FROM `target_sql.customers` as c
JOIN `target_sql.orders` as o on c.customer_id = o.customer_id
GROUP BY
  c.customer_state,
  Year_Month
ORDER BY
  c.customer_state,
  Year_Month;
```

Output Screenshot:

Row	customer_state	Year_Month	number_of_orders
1	AC	2017-01	2
2	AC	2017-02	3
3	AC	2017-03	2
4	AC	2017-04	5
5	AC	2017-05	8
6	AC	2017-06	4
7	AC	2017-07	5
8	AC	2017-08	4
9	AC	2017-09	5
10	AC	2017-10	6

Insights:

The query results can reveal monthly order counts for each state, revealing trends, patterns, or seasonality. This information can help identify states with high order volumes and identify significant changes in order counts. For instance, **SP** has the highest order count every month. Analysing these insights can

help target marketing efforts in states with rising order volumes, identify operational issues in states with falling order volumes, or optimize inventory management based on order trends across states.

3.2 How are the customers distributed across all the states?

Answer Query:

```
SELECT customer_state,
       COUNT(DISTINCT customer_unique_id ) AS number_of_unique_customer,
       ROUND(COUNT(DISTINCT customer_unique_id) / (SELECT COUNT(DISTINCT
customer_unique_id) FROM `target_sql.customers`) * 100,2) AS
percentage_of_customer_distribution
FROM `target_sql.customers`
GROUP BY
  customer_state
ORDER BY
  percentage_of_customer_distribution desc;
```

Output Screenshot:

Row	customer_state	number_of_unique_customer	percentage_of_customer_distribution
1	SP	40302	41.94
2	RJ	12384	12.89
3	MG	11259	11.72
4	RS	5277	5.49
5	PR	4882	5.08
6	SC	3534	3.68
7	BA	3277	3.41
8	DF	2075	2.16
9	ES	1964	2.04
10	GO	1952	2.03

Insights:

The report reveals that the state of SP has the highest number of unique customers, accounting for 41.94% of the total customer base. The next two states with significant customer bases are RJ and MG, contributing 12.89% and 11.72% respectively. States like RS, PR, and SC have a considerable number of customers,

each contributing more than 3%. Conversely, AC, AP, and RR have the least number of customers, each contributing less than 0.1%. This information can be used for strategic planning, targeting marketing campaigns in high-customer-concentrated states or exploring expansion opportunities in fewer-customer states.

#Q4 Impact on Economy: Analyze the money movement by e-commerce by looking at order prices, freight and others.

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

Answer Query:

```
-- Yearwise
WITH payments_2017 AS (
  SELECT SUM(p.payment_value) AS total_payment_2017
  FROM `target_sql.payments` AS p
  JOIN `target_sql.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
),

payments_2018 AS (
  SELECT SUM(p.payment_value) AS total_payment_2018
  FROM `target_sql.payments` AS p
  JOIN `target_sql.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
)

SELECT round(((total_payment_2018 - total_payment_2017) / total_payment_2017) *
100,2) AS percentage_increase
FROM payments_2017, payments_2018

-- Monthwise
WITH payments_2017 AS (
  SELECT EXTRACT(MONTH FROM o.order_purchase_timestamp) AS month,
SUM(p.payment_value) AS total_payment_2017
  FROM `target_sql.payments` AS p
  JOIN `target_sql.orders` AS o
  ON p.order_id = o.order_id
```

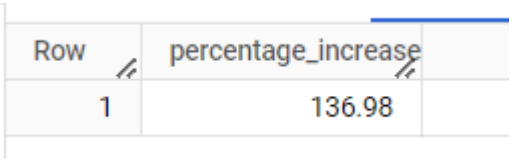
```
  WHERE EXTRACT(YEAR FROM o.order_purchase_timestamp) = 2017
  AND EXTRACT(MONTH FROM o.order_purchase_timestamp) BETWEEN 1 AND 8
  GROUP BY month
),

payments_2018 AS (
  SELECT EXTRACT(MONTH FROM o.order_purchase_timestamp) AS month,
SUM(p.payment_value) AS total_payment_2018
  FROM `target_sql.payments` AS p
  JOIN `target_sql.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
  GROUP BY month
)

SELECT p17.month,round(p17.total_payment_2017,2) as total_payment_2017 ,
round(p18.total_payment_2018,2) as total_payment_2018,
round(((p18.total_payment_2018 - p17.total_payment_2017) / p17.total_payment_2017)
* 100,2) AS percentage_increase
FROM payments_2017 AS p17
JOIN payments_2018 AS p18
ON p17.month = p18.month
ORDER BY
  percentage_increase DESC;
```

Output Screenshot:

-- Yearwise



Row	percentage_increase
1	136.98

-- Monthwise

Row	month	total_payment_2017	total_payment_2018	percentage_increase
1	1	138488.04	1115004.18	705.13
2	2	291908.01	992463.34	239.99
3	4	417788.03	1160785.48	177.84
4	3	449863.6	1159652.12	157.78
5	6	511276.38	1023880.5	100.26
6	5	592918.82	1153982.15	94.63
7	7	592382.92	1066540.75	80.04
8	8	674396.32	1022425.32	51.61

Insights:

The business experienced a significant growth in order volume from 2017 to 2018, with a total cost of orders increasing by 136.98%. The increase was not uniform across all months, with January showing the highest increase at 705.13%. Other months saw substantial increases, with February and April showing 239.99% and 177.84% increases respectively. August saw the lowest increase at 51.61%. The business experienced significant growth in the first half of the year, and further investigation into the factors contributing to this growth could inform future strategies.

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

Answer Query:

```
SELECT c.customer_state,
       ROUND(SUM(p.payment_value),2) AS total_amount ,
       ROUND(AVG(p.payment_value),2) AS avg_amount
FROM `target_sql.payments` AS p
JOIN `target_sql.orders` AS o ON p.order_id = o.order_id
JOIN `target_sql.customers` AS c ON o.customer_id = c.customer_id
GROUP BY c.customer_state
ORDER BY c.customer_state;
```

Output Screenshot:

Row	customer_state	total_amount	avg_amount
1	AC	19680.62	234.29
2	AL	96962.06	227.08
3	AM	27966.93	181.6
4	AP	16262.8	232.33
5	BA	616645.82	170.82
6	CE	279464.03	199.9
7	DF	355141.08	161.13
8	ES	325967.55	154.71
9	GO	350092.31	165.76
10	MA	152523.02	198.86

Insights:

The highest number of unique customers, accounting for 41.94% of the total customer base. The next two states with significant customer bases are RJ and MG, contributing 12.89% and 11.72% respectively.

States like RS, PR, and SC have a considerable number of customers, each contributing more than 3%. Conversely, AC, AP, and RR have the least number of customers, each contributing less than 0.1%.

This information can be used for strategic planning, targeting marketing campaigns in high-customer-concentrated states or exploring expansion opportunities in fewer-customer states.

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

Answer Query:

```
SELECT c.customer_state,
       ROUND(SUM(oi.freight_value),2) AS total_freight ,
       ROUND(AVG(oi.freight_value),2) AS avg_freight
```

```
FROM `target_sql.orders` AS o
JOIN `target_sql.order_items` AS oi ON o.order_id = oi.order_id
JOIN `target_sql.customers` AS c ON o.customer_id = c.customer_id
GROUP BY c.customer_state
ORDER BY c.customer_state;
```

Output Screenshot:

Row	customer_state	total_freight	avg_freight
1	AC	3686.75	40.07
2	AL	15914.59	35.84
3	AM	5478.89	33.21
4	AP	2788.5	34.01
5	BA	100156.68	26.36
6	CE	48351.59	32.71
7	DF	50625.5	21.04
8	ES	49764.6	22.06
9	GO	53114.98	22.77
10	MA	31523.77	38.26

Insights:

- SP has the highest total freight value at 718,723.07, indicating high customer volume.
- RJ and MG have high total freight values, indicating a large volume of orders.
- RR, PB, and RO have high average freight values, indicating high freight cost per order despite smaller volume.
- MS, MT, and PA have average freight values in the mid-range, indicating moderate freight cost per order.
- This information can aid strategic decisions related to pricing, logistics, and customer service.
- Other factors like order volume, product mix, and delivery distance should also be considered.

#Q5 Analysis based on sales, freight and delivery time.

5.1 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:

- **time_to_deliver = order_delivered_customer_date - order_purchase_timestamp**
- **diff_estimated_delivery = order_delivered_customer_date - order_estimated_delivery_date**

Answer Query:

```
SELECT  order_id,
      DATE_DIFF(DATE(order_delivered_customer_date),DATE(order_purchase_timestamp),
DAY) AS delivery_time,
      DATE_DIFF(DATE(order_delivered_customer_date),DATE(order_estimated_delivery_
date), DAY) AS diff_estimated_delivery
FROM  `target_sql.orders`
ORDER BY  order_id;
```

Output Screenshot:

Row	order_id	delivery_time	diff_estimated_delivery
1	00010242fe8c5a6d1ba2dd792...	7	-9
2	00018f77f2f0320c557190d7a1...	16	-3
3	000229ec398224ef6ca0657da...	8	-14
4	00024acbcdf0a6daa1e931b03...	6	-6
5	00042b26cf59d7ce69dfabb4e...	25	-16
6	00048cc3ae777c65dbb7d2a06...	7	-15
7	00054e8431b9d7675808bcb8...	8	-17
8	000576fe39319847cbb9d288c...	5	-16
9	0005a1a1728c9d785b8e2b08...	10	0
10	0005f50442cb953dcd1d21e1f...	2	-19

Insights:

The "Time_to_deliver" column in Query Analysis shows the number of days it takes to deliver an order from the purchase date, while the "diff_estimated_delivery" column indicates the difference between the estimated and actual delivery dates. Analyzing this data helps identify longer-taking orders and compares them with the average timeline to assess delivery efficiency.

Negative values indicate delayed deliveries, while positive values indicate early deliveries. Understanding these variances can improve delivery timelines and logistics, leading to enhanced processes.

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

Answer Query:

```
-- 5.2 Find out the top 5 states with the highest & lowest average freight value.
WITH highest_freight AS (
SELECT c.customer_state,
       ROUND(AVG(oi.freight_value),2) AS total_freight
FROM `target_sql.orders` AS o
JOIN `target_sql.order_items` AS oi ON o.order_id = oi.order_id
JOIN `target_sql.customers` AS c ON o.customer_id = c.customer_id
GROUP BY c.customer_state
ORDER BY total_freight DESC
limit 5 ),

lowest_freight AS (
SELECT c.customer_state,
       ROUND(AVG(oi.freight_value),2) AS total_freight
FROM `target_sql.orders` AS o
JOIN `target_sql.order_items` AS oi ON o.order_id = oi.order_id
JOIN `target_sql.customers` AS c ON o.customer_id = c.customer_id
GROUP BY c.customer_state
ORDER BY total_freight
limit 5)

SELECT CONCAT("HIGH FREIGHT : ", customer_state) AS customer_state, total_freight
FROM highest_freight
UNION ALL
SELECT CONCAT("LOW FREIGHT : ", customer_state)  AS customer_state, total_freight
FROM lowest_freight;
```

Output Screenshot:

Row	customer_state	total_freight
1	HIGH FREIGHT : RR	42.98
2	HIGH FREIGHT : PB	42.72
3	HIGH FREIGHT : RO	41.07
4	HIGH FREIGHT : AC	40.07
5	HIGH FREIGHT : PI	39.15
6	LOW FREIGHT : SP	15.15
7	LOW FREIGHT : PR	20.53
8	LOW FREIGHT : MG	20.63
9	LOW FREIGHT : RJ	20.96
10	LOW FREIGHT : DF	21.04

Insights:

The data from a SQL query reveals that the states with the highest average freight values are RR, PB, RO, AC, and PI, with freight values ranging from 39.15 to 42.98. These states are either far from distribution centers or have difficult terrains, increasing transportation costs.

Conversely, the states with the lowest average freight values are SP, PR, MG, RJ, and DF, with freight values ranging from 15.15 to 21.04. These insights can help businesses understand their logistics costs and strategize their supply chain management. However, these are just insights and should be considered alongside other factors.

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

Answer Query:

```
WITH cte AS (
  SELECT
    customer_state AS state,
    AVG(datetime_diff(order_delivered_customer_date, order_purchase_timestamp, DAY))
AS avg_delivery_time
  FROM
    `target_sql.customers` AS c
  JOIN
    `target_sql.orders` AS o
  ON
```

```
    c.customer_id = o.customer_id
  GROUP BY
    state
),
rankings AS (
  SELECT
    state,
    avg_delivery_time,
    DENSE_RANK() OVER (ORDER BY avg_delivery_time DESC) as rnk_fast,
    DENSE_RANK() OVER (ORDER BY avg_delivery_time ASC) as rnk_slow
  FROM
    cte
)
SELECT
  CASE
    WHEN rnk_fast <= 5 THEN CONCAT('FAST - ', rnk_fast)
    WHEN rnk_slow <= 5 THEN CONCAT('SLOW - ', rnk_slow)
  END AS speed_of_delivery,
  state,
  ROUND(avg_delivery_time, 2) AS Avg_delivery_time
FROM
  rankings
WHERE
  rnk_fast <= 5 OR rnk_slow <= 5
ORDER BY
  speed_of_delivery;
```

Output Screenshot:

Row	speed_of_delivery	state	Avg_delivery_time
1	FAST - 1	RR	28.98
2	FAST - 2	AP	26.73
3	FAST - 3	AM	25.99
4	FAST - 4	AL	24.04
5	FAST - 5	PA	23.32
6	SLOW - 1	SP	8.3
7	SLOW - 2	PR	11.53
8	SLOW - 3	MG	11.54
9	SLOW - 4	DF	12.51
10	SLOW - 5	SC	14.48

Insights:

The query reveals that the states with the fastest average delivery times are RR, AP, AM, AL, and PA, with average delivery times ranging from 23.32 to 28.98 days. These states have

efficient logistics and delivery processes, leading to faster delivery times. Conversely, the states with the slowest average delivery times are SP, PR, MG, DF, and SC, with average delivery times ranging from 8.3 to 14.48 days.

Delivery time variation is significant across states, with the fastest state (RR) having an average delivery time more than three times longer than the slowest state (SP), possibly due to factors like geographical distance, infrastructure, and operational efficiency.

Staying ahead and studying the logistics and delivery processes of states with faster delivery times could help reduce delivery times and improve customer satisfaction.

5.4 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.

Answer Query:

```
SELECT  c.customer_state,
      ROUND(AVG(DATE_DIFF(o.order_estimated_delivery_date,
o.order_delivered_customer_date, DAY)),2) AS diff_estimated_delivery
FROM  `target_sql.orders` AS o
JOIN `target_sql.customers` AS c ON o.customer_id = c.customer_id
GROUP BY c.customer_state
ORDER BY  diff_estimated_delivery DESC
limit 5;
```

Output Screenshot:

Row	customer_state	diff_estimated_delivery
1	AC	19.76
2	RO	19.13
3	AP	18.73
4	AM	18.61
5	RR	16.41

Insights:
The query reveals that the states with the fastest delivery times are AC, RO, AP, AM, and RR, based on the difference between actual and estimated delivery dates. These states have an average delivery efficiency of 19.76 days earlier than the estimated date, followed closely by RO at 19.13 days.

This suggests that the delivery process is highly efficient and ahead of schedule. Faster delivery times can lead to higher customer satisfaction, as customers generally appreciate receiving their orders earlier than expected. The top-performing states can learn from their logistics and delivery processes to improve their own delivery efficiency and customer satisfaction.

#Q6 Alysis based on the payments:

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

Answer Query:

```
SELECT
       FORMAT_TIMESTAMP('%Y-%m', o.order_purchase_timestamp) AS month,
       p.payment_type,
       COUNT(DISTINCT o.order_id) AS order_count
FROM   `target_sql.payments` p
JOIN `target_sql.orders` o ON o.order_id = p.order_id
GROUP BY
       p.payment_type, month
ORDER BY
       month, LOWER(p.payment_type);
```

Output Screenshot:

Row	month	payment_type	order_count
1	2016-09	credit_card	3
2	2016-10	credit_card	253
3	2016-10	debit_card	2
4	2016-10	UPI	63
5	2016-10	voucher	11
6	2016-12	credit_card	1
7	2017-01	credit_card	582
8	2017-01	debit_card	9
9	2017-01	UPI	197
10	2017-01	voucher	33

Insights:

Credit card payments are the most popular payment method, possibly reflecting customer demographics. UPI usage is growing due to increased digital adoption and convenience. Vouchers are used less frequently but consistently, indicating a consistent customer segment. Some payment types are not defined, possibly due to data entry errors or other issues.

The number of orders placed using different payment types is increasing, possibly indicating customer growth or increased sales activity. Overall, credit card payments remain the most popular payment method, indicating a preference for digital payment methods among customers.

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

Answer Query:

```
SELECT payment_installments,
       COUNT(DISTINCT order_id) AS order_count
FROM `target_sql.payments`
WHERE payment_installments <> 0
GROUP BY payment_installments
ORDER BY payment_installments;
```

Output Screenshot:

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

Insights:
The majority of orders (49,060) are placed with a single payment installment, suggesting that most customers prefer to pay in full at the time of purchase.

A decreasing trend in orders as the number of installments increases suggests fewer customers opt for a higher number of installments. However, there are exceptions to this trend, such as more orders with 10 installments (5,315) and more orders with 8 installments (4,253), possibly due to specific promotional offers or financial products.

Installment plans with more than 10 installments are used rarely, with only 18 orders having 24 installments, possibly due to higher interest rates or longer commitment periods.

Recommendations

1. Enhance Regional Marketing & Customer Engagement

- **Focus on High-Volume Regions:** Since states like SP, RJ, and MG have the highest number of orders and unique customers, tailor marketing campaigns and localized promotions for these areas.
- **Target Under-Served Areas:** In regions with low customer density (e.g., AC, AP, RR), consider targeted outreach or localized incentives to boost customer engagement and market penetration.

2. Leverage Seasonal Trends

- **Optimize Inventory for Peak Periods:** The data shows a marked increase in orders during holiday seasons (e.g., November) and promotional events (e.g., FIFA World Cup). Increase stock and staffing ahead of these periods to meet higher demand.
- **Plan Promotions Strategically:** Use the observed seasonality to schedule special promotions or discounts during peak months, ensuring that marketing efforts are aligned with customer purchasing behavior.

3. Streamline Logistics & Delivery Operations

- **Analyze Delivery Time Variations:** Significant differences in average delivery times across states suggest room for improvement in logistics. Investigate the processes in states with the fastest deliveries (e.g., SP, PR, MG) and implement similar best practices in slower regions.
- **Address Freight Cost Disparities:** With states like RR, PB, and RO showing higher average freight values, consider negotiating better shipping rates or optimizing routes to reduce costs, which can also enhance customer satisfaction through more predictable delivery times.

4. Optimize Payment Processes

- **Capitalize on Preferred Payment Methods:** Since credit cards are the most popular and UPI usage is growing, invest in enhancing the digital payment experience. Ensure that the payment gateway is secure and user-friendly.
- **Review Installment Options:** The data shows that most customers prefer single-installment payments, but a noticeable number also opt for multi-installment plans. Consider introducing flexible payment options or promotions that make installment payments more attractive, especially during high-value promotions.

5. Use Data-Driven Insights for Continuous Improvement

- **Monitor and React to Trends:** Set up regular reviews of key performance indicators (KPIs) related to order volumes, freight values, and delivery times. Use these insights to make agile adjustments to marketing, inventory, and logistics strategies.
- **Integrate Customer Feedback:** While quantitative insights are essential, integrating qualitative feedback (such as customer reviews) can provide a fuller picture of the guest experience, helping refine both product offerings and service quality.

Implementing these recommendations could help Target not only sustain its growth in Brazil but also improve operational efficiency and customer satisfaction.