Business Case: Target SQL

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 `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

Actionable Insights:

- The dataset includes `customer_id` and `customer_unique_id` and both are of string data type.
- The dataset includes `customer_zip_code_prefix`, ` customer_city` and `customer_state`.

Recommendations:

- Investigate the uniqueness of customer_id and customer_unique_id pairs. Duplicate entries, missing or incomplete values might indicate potential data quality issues that need cleaning or further investigation.
- Analyze zip code prefixes, customer city and state to understand customer distribution and regional preferences. This could aid in creating personalized marketing campaigns targeting specific regions.

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

Query:

```
SELECT MIN(order_purchase_timestamp) AS start_date,
MAX(order_purchase_timestamp) AS end_date
FROM `target.orders`;
```

Output:

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

Actionable Insights:
Orders were placed between September 2016 and October 2018.

Recommendations:
Assess the performance over this period to identify trends, growth patterns and seasonal peaks. We can use these insights to guide our future marketing and inventory planning.

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

Query:

```
SELECT COUNT(DISTINCT customer_city) AS num_cities,
COUNT(DISTINCT customer_state) AS num_states
FROM `target.customers`;
```

OR

```
SELECT COUNT(DISTINCT customer_city) AS num_cities,
COUNT(DISTINCT customer_state) AS num_states
FROM `target.customers` c JOIN `target.orders` o
ON c.customer_id=o.customer_id
WHERE order_purchase_timestamp BETWEEN '2016-09-04' AND '2018-10-17';
```

Output:

Row	num_cities	num_states
1	4119	27

Actionable Insights:
Orders are spread across 4119 cities and 27 states.

Recommendations:
The above data can be compared with the overall counts for the state (27) and cities (8011).

- Evaluate state-wise order distribution. Identify high-performing states for potential expansion and target underrepresented states for growth.
- Focus on marketing efforts and research for the cities with few or no orders and plan our campaigns accordingly.

Note:
The first query provides an overall count of cities and states without considering any time constraints.

- Doesn't involve any joins, making it faster to execute.
- Suitable when we are confident that the dataset is within the specified time range, for ex. the target dataset on which we are currently working.

The second query considers the specified time range and provides counts only for customers who placed orders within that range, which will be helpful for further analysis if we want to change the range.

2. **In-depth Exploration:**

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

Query 1:

```
SELECT EXTRACT(YEAR FROM order_purchase_timestamp) AS order_year,
COUNT(DISTINCT order_id) AS num_orders
FROM `target.orders`
GROUP BY order_year
ORDER BY order_year;
```

Output:

Row	order_year	num_orders
1	2016	329
2	2017	45101
3	2018	54011

Query 2:

```
SELECT FORMAT_TIMESTAMP('%Y-%m', order_purchase_timestamp) AS year_month,
COUNT(DISTINCT order_id) AS num_orders
FROM `target.orders`
GROUP BY year_month
ORDER BY year_month;
```

Output:

Row	year_month	num_orders
1	2016-09	4
2	2016-10	324
3	2016-12	1
4	2017-01	800
5	2017-02	1780
6	2017-03	2682
7	2017-04	2404
8	2017-05	3700
9	2017-06	3245
10	2017-07	4026

Actionable Insights:
A notable upward trend in orders from 2016 to 2018, more specifically from Jan 2017 to Aug 2018.

Recommendations:
Capitalize on the growth pattern by investing in scaling strategies, such as inventory management, optimized logistics and expanded marketing.

Assumption:
Note that the orders were placed between Sep 2016 and Oct 2018, so we only have about two years of data. So, the first query is not sufficient to identify the exact growing trend over the years. Therefore, for in-depth exploration, I used the second query.

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

Query:

```
SELECT FORMAT_TIMESTAMP('%Y-%m', order_purchase_timestamp) AS year_month,
COUNT(DISTINCT order_id) AS num_orders
FROM `target.orders`
GROUP BY year_month
ORDER BY year_month;
```

Output:

Row	year_month	num_orders
1	2016-09	4
2	2016-10	324
3	2016-12	1
4	2017-01	800
5	2017-02	1780
6	2017-03	2682
7	2017-04	2404
8	2017-05	3700
9	2017-06	3245
10	2017-07	4026

Actionable Insights:

- Gradual Growth: Overall, there's a trend of gradual growth in order volumes from January 2017 to August 2018.
- Peak Shopping Season: Months like November, December and January consistently exhibit higher order volumes, likely due to holiday shopping, new-year and post-holiday sales.
- Mid-Year Peaks: Mid-year months, particularly July and August, also show increased order counts, possibly influenced by seasonal trends or events.
- Low Season Months: September and October have comparatively lower order counts, suggesting a post-holiday lull.

Recommendations:

- Align marketing with peak months, capitalizing on high-demand periods.
- Optimize inventory for demand spikes, ensuring adequate stock levels.
- Utilize lull months for inventory review, system upgrades and training.
- Engage customers by understanding seasonal behavior and tailoring personalized offers.

Assumption:

The fluctuation in orders in the fourth quarter of 2017 might be due to the business's initial phase and gradual growth, as it started. Order drop in Sep and Oct 2018 may be due to data limitations, incompleteness or potential truncation. So, we analyze monthly trends by excluding these outliers to avoid skewing seasonality interpretation.

3. 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

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'
    ELSE 'Night'
  END AS time_of_day,
  COUNT(DISTINCT order_id) AS num_orders
FROM `target.orders`
GROUP BY time_of_day
ORDER BY num_orders DESC;
```

Output:

Row	time_of_day	num_orders
1	Afternoon	38135
2	Night	28331
3	Morning	27733
4	Dawn	5242

Actionable Insights:

- Afternoon Dominance: Brazilian customers predominantly place orders in the afternoon.
- Consistent Mornings and Nights: Morning and night orders are relatively balanced, suggesting stable early and late shopping habits.
- Dawn's Lowest Count: Dawn witnesses the fewest orders, possibly due to its less conducive shopping time.

Recommendations:

- Capitalize on the afternoon peak by scheduling promotions and campaigns during this time for higher visibility and engagement.
- Consider targeting early morning customers with exclusive offers to drive sales, this might attract early risers to shop during dawn.
- Extend customer support for night orders to enhance satisfaction, addressing queries promptly and improving the shopping experience during evenings.

Assumption:
Ordering behavior is influenced by local customs, work patterns, online shopping trends and seasons. Afternoon is popular due to convenience, evenings suit leisure, weekends allow more time and special occasions might alter timing.

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 c.customer_state,
EXTRACT(MONTH FROM o.order_purchase_timestamp) AS order_month,
COUNT(DISTINCT o.order_id) AS num_orders
FROM `target.orders` o RIGHT JOIN `target.customers` c
ON o.customer_id=c.customer_id
GROUP BY c.customer_state, order_month
ORDER BY c.customer_state, order_month;
```

Output:

Row	customer_state	order_month	num_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
11	AC	11	5
12	AC	12	5
13	AL	1	39
14	AL	2	39

Actionable Insights:

- Seasonal and State Variation: Different states exhibit diverse order patterns and display monthly fluctuations tied to consumer behaviors and local preferences.
- Key State Dynamics:
    - SP consistently leads in orders, emphasizing its major market role.
    - Emerging growth in states like RJ and MG indicates untapped e-commerce potential.
- Population Influence: States with larger populations exhibit higher e-commerce order counts across the year, reflecting a larger consumer base and market demand.
- Geographical Challenges: States with lower order counts (RR, AP, AC) face geographical barriers, limited internet access, and less urbanization, impacting e-commerce adoption.

Recommendations:

- Tailor strategies to state-specific patterns, preferences and challenges.
- Strategically allocate resources based on population and economic significance.
- Enhance e-commerce infrastructure in major economic centers.
- Address challenges in states with geographical barriers through innovative solutions.
- Focus on improving digital connectivity and literacy in underserved areas.

Assumption:

The dataset accurately represents customer distribution across states. The analysis is based on available data and factors external to the dataset are not considered.
Note that the output has 322 rows, which should be (27 states * 12 months = 324). Further analysis led me to the conclusion that no orders are placed for the state RR during the months of August and December.

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

Query:

```
SELECT customer_state,
COUNT(DISTINCT customer_unique_id) AS num_of_customers
FROM `target.customers`
GROUP BY customer_state
ORDER BY num_of_customers desc;
```

Output:

Row	customer_state	num_of_customers
1	SP	40302
2	RJ	12384
3	MG	11259
4	RS	5277
5	PR	4882
6	SC	3534
7	BA	3277
8	DF	2075
9	ES	1964
10	GO	1952

Actionable Insights:

- Concentration of Customers: States like SP, RJ and MG exhibit higher customer counts, indicating a strong e-commerce presence in densely populated and economically active regions.
- Regional Variations: Customer distribution varies widely across states. Some states, like DF, ES and GO, have significant customer bases, while others have comparatively lower numbers.
- Geographical Impact: Less populated and geographically remote states such as AC, AP, RP and TO have lower customer counts, which could be attributed to factors like connectivity challenges.

Recommendations:

- Focus marketing and optimize inventory in states like SP, RJ, MG. Tailor campaigns to engage and retain customers effectively.
- Expand in Growing States: Allocate resources for emerging markets. Forge local partnerships for deeper market penetration.
- Bridge Connectivity Gaps: Collaborate for digital enhancements in remote states (AC, AP, RR, TO). Provide incentives for online shopping to mitigate connectivity challenges.

Assumption:

The dataset accurately represents customer distribution across states. The analysis is based on available data and factors external to the dataset are not considered.

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 yearly_total_costs AS (
SELECT
EXTRACT(YEAR FROM o.order_purchase_timestamp) AS year,
ROUND(SUM(p.payment_value),2) AS total_payment
FROM `target.orders` o JOIN `target.payments` p ON o.order_id = p.order_id
WHERE EXTRACT(YEAR FROM o.order_purchase_timestamp) IN (2017, 2018)
AND EXTRACT(MONTH FROM o.order_purchase_timestamp) BETWEEN 1 AND 8
GROUP BY year
ORDER BY year
)

SELECT (SELECT total_payment FROM yearly_total_costs WHERE year=2017) AS
`total_cost_2017`,
(SELECT total_payment FROM yearly_total_costs WHERE year=2018) AS `total_cost_2018`,
ROUND(((SELECT total_payment FROM yearly_total_costs WHERE year=2018)-(SELECT
total_payment FROM yearly_total_costs WHERE year=2017))/(SELECT total_payment FROM
yearly_total_costs WHERE year=2017)*100,2) AS percentage_increase;
```

Output:

Row	total_cost_2017	total_cost_2018	percentage_increase
1	3669022.12	8694733.84	136.98

Actionable Insights:
The percentage increase in the cost of orders from Jan to Aug 2017 to the same period in 2018 is 136.98%.

Recommendations:

- E-commerce businesses should leverage this significant increase to further enhance their offerings and customer experience.
- Conduct a detailed analysis to identify the factors contributing to this growth, such as marketing strategies, product diversification or improved customer engagement.
- Prepare strategies to sustain and manage future growth, ensuring scalability and efficiency in operations.

Assumption:
The analysis assumes accurate and complete data for order costs in both 2017 and 2018. The percentage increase is based on the local currency used for order costs, without considering inflation or currency fluctuations.

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

Query:

```
SELECT c.customer_state,
ROUND(SUM(p.price),2) AS total_price,
ROUND(AVG(p.price),2) AS average_price
FROM `target.customers` c JOIN `target.orders` o ON c.customer_id=o.customer_id
JOIN `target.order_items` p ON o.order_id=p.order_id
GROUP BY c.customer_state
ORDER BY total_price desc, average_price desc;
```

Output:

Row	customer_state	total_price	average_price
1	SP	5202955.05	109.65
2	RJ	1824092.67	125.12
3	MG	1585308.03	120.75
4	RS	750304.02	120.34
5	PR	683083.76	119.0
6	SC	520553.34	124.65
7	BA	511349.99	134.6
8	DF	302603.94	125.77
9	GO	294591.95	126.27
10	ES	275037.31	121.91
11	PE	262788.03	145.51
12	CE	227254.71	153.76

Actionable Insights:

- SP has the highest total order price due to its large population and economic activity, but its lower average suggests a wide range of order values.
- RR's lowest total indicates lower market engagement, while PB's highest average implies a concentration of higher-priced orders.
- Other states exhibit varying total and average order prices, reflecting diverse consumer behaviors, market dynamics and economic conditions.

Recommendations:

- Allocate resources to capitalize on high-performing states by tailoring marketing and product offerings.
- Focus on states with lower total prices by introducing attractive promotions or targeting specific customer segments.
- Target states with higher average prices for premium offerings and personalized experiences, while working to improve average prices in states with potential.

Assumption:

The analysis assumes accurate and complete order price data for all states. Differences in total and average prices could result from varying economic affluence and consumer demographics.

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

Query:

```
SELECT c.customer_state,
ROUND(SUM(p.freight_value),2) AS total_freight,
ROUND(AVG(p.freight_value),2) AS average_freight
FROM `target.customers` c JOIN `target.orders` o ON c.customer_id=o.customer_id
JOIN `target.order_items` p ON o.order_id=p.order_id
GROUP BY c.customer_state
ORDER BY total_freight desc, average_freight desc;
```

Output:

Row	customer_state	total_freight	average_freight
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
11	ES	49764.6	22.06
12	CE	48351.59	32.71

Actionable Insights:

- There are significant variations in total and average order freight costs across states.
- SP has the highest total freight but the lowest average, while RR has the lowest total freight but the highest average.
- RR, AP, AC and AM show higher average freight costs, potentially indicating unique shipping challenges or preferences in these states.

Recommendations:

- Analyze states with high average freight costs to identify potential inefficiencies in shipping processes and explore options for cost reduction.
- States with lower average freight costs are more appealing to customers. Consider leveraging this as a marketing point to attract more orders.
- Explore partnerships with local logistics providers to potentially reduce shipping costs and improve service quality in specific regions.

Assumption:

The analysis assumes accurate and complete data for order freight costs in each state. The analysis assumes that the freight costs are in the same currency across all states.

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

Query:

```
SELECT order_id,
DATETIME_DIFF(order_delivered_customer_date,order_purchase_timestamp,DAY)
time_to_deliver,
DATETIME_DIFF(order_estimated_delivery_date,order_delivered_customer_date,DAY)
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
11	66057d37308e787052a32828...	38	-6
12	19135c945c554eebfd7576c73...	36	-2

Actionable Insights:

- Delivery time distribution is right-skewed, ranging from 0 to 209 days. Predominantly, 97.55% of orders are delivered within 30 days, with a peak at 7 days. Around 3.68% fall in the 30-50 days range, while the remaining 0.77% extend beyond.
- The `diff_estimated_delivery` varies from -188 to 146 days. Around 97% of orders have a difference between estimated and actual delivery dates within -20 to 30 days, peaking at 7 days, while 3% fall outside this range.

Recommendations:

- Efficient Delivery Optimization: Focus on further improving delivery times for orders beyond 30 days to enhance customer satisfaction and loyalty.
- Estimation Refinement: Fine-tune the estimation process to minimize the variance between estimated and actual delivery dates for the majority of orders.
- Customer Communication: For orders with delivery times beyond 30 days or discrepancies in estimation, communicate proactively with customers to manage expectations.

Assumptions:

The analysis assumes accurate time tracking for order purchase, delivery, and estimated delivery. It also presupposes consistent data across orders and reliable estimation of delivery times by the business.

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

Query:

```
WITH customer_average_freight AS (
  SELECT c.customer_state,
  ROUND(AVG(p.freight_value),2) AS average_freight
  FROM `target.customers` c JOIN `target.orders` o ON c.customer_id=o.customer_id
  JOIN `target.order_items` p ON o.order_id=p.order_id
  GROUP BY c.customer_state
)

(SELECT *
FROM customer_average_freight
ORDER BY average_freight
LIMIT 5)
UNION ALL
(SELECT *
FROM customer_average_freight
ORDER BY average_freight DESC
LIMIT 5)
ORDER BY average_freight;
```

Output:

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

Actionable Insights:

- States like SP, PR, MG, RJ and DF have the lowest average freight values, indicating efficient shipping or strong logistics networks.
- States such as AC, RO, PB, RR and PI have notably higher average freight values, potentially due to remote locations or limited connectivity.

Recommendations:

- For states with lower average freight values, further optimize logistics and operations to maintain cost-effectiveness.
- Focus on states with high average freight values. Negotiate better shipping rates, address challenges, and tailor solutions for geographical efficiency.
- Investigate why SP has the lowest average freight value. Identify cost-effective shipping solutions that other states could adopt.

Assumptions:

The analysis assumes accurate freight data. Freight values vary due to factors like distance, shipping partners, and regional logistics challenges.

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

Query:

```
WITH customer_average_delivery_time AS (
  SELECT c.customer_state,
ROUND(AVG(DATETIME_DIFF(order_delivered_customer_date,order_purchase_timestamp,DAY)),2)
average_delivery_time
  FROM `target.customers` c JOIN `target.orders` o ON c.customer_id=o.customer_id
  GROUP BY c.customer_state
)

(SELECT *
FROM customer_average_delivery_time
ORDER BY average_delivery_time
LIMIT 5)
UNION ALL
(SELECT *
FROM customer_average_delivery_time
ORDER BY average_delivery_time desc
LIMIT 5)
ORDER BY average_delivery_time;
```

Output:

Row	customer_state	average_delivery_time
1	SP	8.3
2	PR	11.53
3	MG	11.54
4	DF	12.51
5	SC	14.48
6	PA	23.32
7	AL	24.04
8	AM	25.99
9	AP	26.73
10	RR	28.98

Actionable Insights:

- The top 5 states with the lowest average delivery times are SP, PR, MG, DF and SC, with delivery times ranging from 8 to 15 days.
- The bottom 5 states with the highest average delivery times are PA, AL, AM, AP and RR, with delivery times ranging from 23 to 29 days.

Recommendations:

- States with longer delivery times should enhance logistics for quicker transit and improved customer satisfaction.
- Invest in better transportation and distribution networks to notably impact delivery times, especially in states with extended averages.
- Strategically place regional warehouses to minimize distances, leading to faster deliveries.
- Implement real-time package tracking for transparency and reduced concerns.
- Collaborate with trusted logistics partners for optimized routes and streamlined delivery processes.

Assumptions:

The analysis considers uniform measurement and assumes consistent shipping methods and demand. External factors like weather and regulations are excluded from this data-driven assessment.

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:

```
SELECT customer_state,
ROUND(AVG(DATETIME_DIFF(order_estimated_delivery_date,order_delivered_customer_date,DAY))
,2) AS est_del_diff
FROM `target.customers` c JOIN `target.orders` o ON c.customer_id=o.customer_id
GROUP BY customer_state
ORDER BY est_del_diff desc
LIMIT 5;
```

Output:

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

Actionable Insights:

- The top 5 states with the fastest order delivery compared to the estimated date of delivery are AC, RO, AP, AM and RR, with differences ranging from 17 to 20 days.
- Geographically challenged states with maximum delivery times unexpectedly show quicker deliveries as compared to estimated deliveries, likely due to optimized strategies.

Recommendations:

- Capitalize on local knowledge to optimize delivery routes, considering unique geographical characteristics.
- Adjust estimated delivery times to reflect the demonstrated efficiency in these challenging regions.
- Implement advanced navigation and tracking systems to navigate complex terrains more effectively.
- Disseminate successful strategies from these states to others facing similar geographical challenges.

Assumptions:

The analysis assumes accurate time tracking for order delivery and estimated delivery. External factors like weather and regulations are excluded from this data-driven assessment.

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

Query:

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

Output:

Row	month	payment_type	num_orders
1	01	UPI	1715
2	01	credit_card	6093
3	01	debit_card	118
4	01	voucher	337
5	02	UPI	1723
6	02	credit_card	6582
7	02	debit_card	82
8	02	voucher	288
9	03	UPI	1942
10	03	credit_card	7682

Actionable Insights:

- Orders placed with different payment types fluctuate across months.
- Credit card is the primary payment choice followed by UPI, while debit card and vouchers are less frequent.
- There are 3 orders with an unspecified payment type recorded in the months of August and September.

Recommendations:

- Encourage diverse payment options to cater to varying customer preferences.
- Ensure secure payment gateways for all payment types to build trust among customers.
- Offer incentives for using certain payment methods to promote customer adoption.
- Boost promotions for debit card usage to increase its adoption rate.

Assumptions:
The analysis relies on precise data for order payment types. Regional payment preferences, external factors and marketing impacts aren't included.

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 num_orders
FROM `target.payments`
WHERE payment_installments>=1 and payment_value>0
GROUP BY payment_installments
ORDER BY payment_installments;
```

Output:

Row	payment_installments	num_orders
1	1	49057
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

Actionable Insights:

- Most orders are concentrated within a lower range of payment installments.
- There's a gradual decline in the number of orders as the number of payment installments increases beyond 10.
- A few orders have higher installments, with a notable drop in frequency.

Recommendations:

- Promote flexible payment options for varied preferences.
- Communicate installment plans' benefits transparently.
- Investigate and manage reasons behind uncommonly high installment counts for better customer satisfaction.

Assumption:

Payments are accurately recorded and reflect successful payment installments. No external factors affecting payment choices are considered in this analysis.
Orders with 0 payment installments are excluded to focus on the analysis's targeted context and align with the question's intent.