Business Case-Target Sql

Sunku Sai Purnima

DSML-OCT23-MWF

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.

```
SELECT *
FROM `target.INFORMATION_SCHEMA.COLUMNS`
WHERE table_name='customers';
```

Insights: In this table, customer general details such as their location are stored, each allocated a unique ID for those who have placed orders through the site.

Query results

SAVE RESULTS

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

Row	table_catalog	table_schema	table_name	column_name	ordinal_position	is_nullable	data_type	column_default
1	ecommerce-406204	target	customers	customer_id	1	YES	STRING	NULL
2	ecommerce-406204	target	customers	customer_unique_id	2	YES	STRING	NULL
3	ecommerce-406204	target	customers	customer_zip_code_prefix	3	YES	INT64	NULL
4	ecommerce-406204	target	customers	customer_city	4	YES	STRING	NULL
5	ecommerce-406204	target	customers	customer_state	5	YES	STRING	NULL

#1B. Get the time range between which the orders were placed.

```
SELECT
   FORMAT_DATETIME("%c",MIN(order_purchase_timestamp)) AS `first_order`,
   FORMAT_DATETIME("%c", MAX(order_purchase_timestamp)) AS `last_order`,
   DATETIME_DIFF(
   MAX(order_purchase_timestamp),MIN(order_purchase_timestamp), DAY) as `number_of_days`
FROM `target.orders` ;
```

Insights:
The dataset comprises 772 days of data from the Target E-commerce organization, spanning from September 2016 to October 2018.

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

Row	first_order	last_order	number_of_days
1	Sun Sep 4 21:15:19 2016	Wed Oct 17 17:30:18 2018	772

#1 C.Count the Cities & States of customers who ordered during the given period.

```
SELECT
   COUNT(DISTINCT c.customer_city) AS `city_count`,
   COUNT(DISTINCT c.customer_state) AS `state_count`
FROM `target.orders` o
JOIN `target.customers` c
ON o.customer_id=c.customer_id;
```

Insights: The dataset contains information pertaining to "Target" Ecommerce business spanning from September 2016 to October 2018, covering a duration of 772 days. Orders were placed by customers residing in nearly 4,119 cities across 27 different states.

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

Row	city_count	state_count
1	4119	27

2.In-depth Exploration:
#2A.Is there a growing trend in the no. of orders placed over the past years?

```
SELECT
  EXTRACT(YEAR FROM order_purchase_timestamp) `purchases_year`,
  COUNT(order_id) AS `no_of_orders`
FROM `target.orders`
GROUP BY purchases_year
ORDER BY purchases_year;
```

Insights:
- Based on this data, it's evident that there is a growth trend in the number of orders placed each year:
- There was a significant increase from 2016 (329 orders) to 2017 (45,101 orders).
- Another increase is observed from 2017 to 2018 (54,011 orders), indicating continued growth in the number of orders placed.

Therefore, it appears that there is a noticeable upward trend in the number of orders placed over the past years, showing an increase from year to year.

JOB INFORMATION | RESULTS | CHART | PREVIEW | JSON

Row	purchases_year	no_of_orders
1	2016	329
2	2017	45101
3	2018	54011

```sql
#2B.Can we see some kind of monthly seasonality in terms of the no. of orders being placed?
SELECT
    -- EXTRACT(MONTH FROM order_purchase_timestamp) AS `month`,
   FORMAT_DATETIME("%B", order_purchase_timestamp) AS `month`,
   COUNT(*) AS `orders_count_month`
FROM `target.orders`
GROUP BY month
ORDER BY orders_count_month DESC;#Ordering by orders count gives highest orders placed months..
```

Insights:

Peak Months: August, May, July, and March remain as the months with higher order counts compared to other months.

Trend: There seems to be a pattern where the summer months (May, June, July, August) have consistently higher order counts compared to the rest of the year. March, being the start of spring in some regions, also shows relatively higher order counts.

Lower Order Counts: September, October, November, and December exhibit lower order counts compared to the preceding months.

December Dip: December notably sees a decrease in order counts, which could be attributed to holiday-related factors such as Christmas and New Year preparations or changes in consumer spending behavior.

Based on this data, there appears to be some monthly seasonality in the number of orders placed, with certain months consistently showing higher order counts compared to others, likely influenced by factors such as seasonality, holidays, or consumer behavior.

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

Row	month	orders_count_month
1	August	10843
2	May	10573
3	July	10318
4	March	9893
5	June	9412
6	April	9343
7	February	8508
8	January	8069
9	November	7544
10	December	5674
11	October	4959
12	September	4305

Recommendations:

By leveraging the insights gained from analyzing monthly order counts, businesses can optimize their strategies, marketing efforts, and inventory management to maximize sales during peak months while maintaining consistent performance during slower periods. Adjusting strategies based on seasonal trends can help businesses better meet customer demand and improve overall performance throughout the year.

#2C.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
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 `num_of_orders`
FROM `target.orders`
GROUP BY time_of_day
ORDER BY num_of_orders DESC;
```

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

Insights:

From this data, we can make the following observations:

Order Distribution by Time of Day: The majority of orders are placed during the Afternoon, followed by Night and Morning, while Dawn has the fewest orders.

Peak Ordering Times: Afternoon seems to be the peak time for orders, with the highest number of orders compared to other times of the day.

Recommendations:

Based on this analysis, here are some potential recommendations:

Optimizing Resources: Allocate more resources, such as customer service representatives or logistics support, during peak times (especially Afternoon) to manage the higher volume of orders efficiently.

Marketing and Promotions: Concentrate marketing efforts or run promotions during times when order counts are typically lower, like Dawn, to potentially stimulate more orders during these periods.

Enhancing Dawn Orders: Investigate reasons behind lower order counts during Dawn. Consider implementing strategies to encourage more orders during this time, such as offering exclusive early-morning deals or targeted marketing campaigns.

```
## 3.Evolution of E-commerce orders in the Brazil region:
#3A.Get the month on month no. of orders placed in each state.
SELECT
EXTRACT(YEAR FROM order_purchase_timestamp) AS `year`,
-- FORMAT_DATETIME("%B", order_purchase_timestamp) AS `month`,
EXTRACT(MONTH FROM order_purchase_timestamp) AS `month`,
c.customer_state,
COUNT(*) AS `num_of_orders`
FROM `target.orders` o
JOIN `target.customers` c
ON o.customer_id=c.customer_id
GROUP BY year,month,c.customer_state
ORDER BY year,month
```

Row	year	month	customer_state	num_of_orders
1	2016	9	RR	1
2	2016	9	RS	1
3	2016	9	SP	2
4	2016	10	SP	113
5	2016	10	RS	24
6	2016	10	BA	4
7	2016	10	PR	19
8	2016	10	RJ	56
9	2016	10	RN	4
10	2016	10	MT	3
11	2016	10	PB	1

Insights:

This analysis demonstrates the month-on-month variation in the number of orders placed in different states. Some states show consistent order placement in specific months, while others exhibit sporadic or no orders during certain periods. Understanding these variations can help in adjusting marketing strategies, inventory management, and resource allocation to optimize sales in specific regions during particular times.

We can visualize the analysis better in chart/graphs using any visualization tool.

#3B.How are the customers distributed across all the states?

```
SELECT
   c.customer_state,
   COUNT(c.customer_unique_id) AS `customer_count`
FROM `target.customers` c
GROUP BY c.customer_state
ORDER BY customer_count DESC;
```

Row	customer_state	customer_count
1	SP	41746
2	RJ	12852
3	MG	11635
4	RS	5466
5	PR	5045
6	SC	3637
7	BA	3380
8	DF	2140
9	ES	2033
10	GO	2020
11	PE	1652
12	CE	1336
13	PA	975
14	MT	907
15	MA	747

Insights:

Based on this data:

- **Higher Customer Counts:** States like SP,RJ and MG have relatively higher customer counts.
- **Moderate Customer Counts:** RS, PR, SC, BA, and SE have moderate customer counts.
- **Lower Customer Counts:** RR,AP,AM have comparatively lower customer counts.

A visualization, such as a bar chart or a map, would better illustrate the proportional distribution of customers across states. It could help in easily identifying states with higher customer concentrations compared to those with lower customer bases.

##4.Impact on Economy: Analyze the money movement by e-commerce by looking at order prices, freight and others.
#4A.Get the % increase in the cost of orders from year 2017 to 2018 (include months between Jan to Aug only).

```
WITH
    cte1 AS (
SELECT
    *
FROM `target.orders` o
JOIN `target.payments` p
ON o.order_id=p.order_id
WHERE EXTRACT(YEAR FROM order_purchase_timestamp) BETWEEN 2017 AND 2018 AND
EXTRACT(MONTH FROM order_purchase_timestamp) BETWEEN 1 AND 8),

    cte2 AS (
SELECT
    EXTRACT(YEAR FROM order_purchase_timestamp) AS `year`,
    ROUND(SUM(payment_value)) AS `cost`
FROM cte1
GROUP BY year
)

SELECT
   *,
   IFNULL(LAG(cost,1) OVER(ORDER BY YEAR),cost) AS `previous_cost`,
   IFNULL(ROUND(((cost-LAG(cost)OVER(ORDER BY YEAR))/
   LAG(cost)OVER(ORDER BY YEAR)*100),2),0) AS `percentage_increase`
FROM cte2
ORDER BY YEAR;
```

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

Row	year	cost	previous_cost	percentage_increase
1	2017	3669022.0	3669022.0	0.0
2	2018	8694734.0	3669022.0	136.98

Insights from Order Prices Analysis:

Revenue Generation: Order prices contribute to the revenue generated by the e-commerce platform. Higher order prices indicate potential revenue growth and customer spending.

Product Pricing Strategies: Understanding the distribution of order prices helps in evaluating the effectiveness of pricing strategies. It helps in identifying popular price ranges and optimizing pricing strategies for different products or services.

Customer Behaviour: Analysis of order prices can reveal customer preferences and buying patterns. It helps in identifying which price points attract more customers or generate higher sales.

From the above result the percentage increase in the cost of orders from January to August between 2017 and 2018 is approximately 136.98%.

Recommendations:

A significant increase in costs (approximately 137.09%) from January to August between 2017 and 2018 warrants a strategic approach to manage and optimize expenses, efficiency improvement, supplier negotiations to handle orders etc.

```sql
#4B. Calculate the Total & Average value of order price for each state.
SELECT
   c.customer_state,
   ROUND(SUM(oi.price),2) AS `Total_value` ,
   ROUND(SUM(oi.price)/COUNT(DISTINCT oi.order_id),2) AS `avg_per_state`
FROM `target.orders` o
JOIN `target.order_items` oi
ON o.order_id=oi.order_id
JOIN `target.customers` c
ON o.customer_id=c.customer_id
GROUP BY c.customer_state
ORDER BY Total_value DESC,avg_per_state DESC
```

Row	customer_state	Total_value	avg_per_state
1	SP	5202955.05	125.75
2	RJ	1824092.67	142.93
3	MG	1585308.03	137.33
4	RS	750304.02	138.13
5	PR	683083.76	136.67
6	SC	520553.34	144.12
7	BA	511349.99	152.28
8	DF	302603.94	142.4
9	GO	294591.95	146.78
10	ES	275037.31	135.82
11	PE	262788.03	159.46
12	CE	227254.71	171.25
13	PA	178947.81	184.48

Analysis of Total Value and Average Order Value per State:

Top Performers by Total Value:

- **SP :** Leads significantly in total order value , indicating a robust market presence and high sales volume.

•Variation in Average Order Values:

- **PB :** Continues to exhibit the highest average order value at $216.67 per order, maintaining a consistently high average across different datasets.

•Observations on Other States:

- **RJ and MG:** Display substantial total order values, indicating strong market potential in these regions.
- **RR , AP , and AC :** Maintain lower total order values and moderate to high average order values, showcasing differing market behaviours.

Recommendations:

Capitalizing on High-Value Markets: Concentrate efforts on states like SP, RJ, and MG to sustain and expand market share due to their high total order values.

Average Order Value Enhancement: Implement strategies to boost average order values in states with lower averages, maximizing revenue per order.

Tailored Regional Approaches: Customize marketing strategies and product assortments based on regional preferences to drive performance in each state.

Enhanced Customer Engagement: Develop customer-centric approaches in states with lower total values to encourage repeat purchases and increase order frequency.

These recommendations aim to refine sales strategies, elevate average order values, and optimize revenue by understanding the diverse performance across different states in the e-commerce dataset.

```
#C.Calculate the Total & Average value of order freight for each state.
SELECT
   c.customer_state,
   ROUND(SUM(oi.freight_value),2) AS `Total_value` ,
    ROUND(SUM(oi.freight_value)/COUNT(DISTINCT oi.order_id),2) AS `avg_per_state`
FROM `target.orders` o
JOIN `target.order_items` oi
ON o.order_id=oi.order_id
JOIN `target.customers` c
ON o.customer_id=c.customer_id
GROUP BY c.customer_state
ORDER BY Total_value DESC,avg_per_state DESC
```

Row	customer_state	Total_value	avg_per_state
1	SP	718723.07	17.37
2	RJ	305589.31	23.95
3	MG	270853.46	23.46
4	RS	135522.74	24.95
5	PR	117851.68	23.58
6	BA	100156.68	29.83
7	SC	89660.26	24.82
8	PE	59449.66	36.07
9	GO	53114.98	26.46
10	DF	50625.5	23.82
11	ES	49764.6	24.58
12	CE	48351.59	36.44

Analysis of Total Value and Average Order Value per State:

High Total Order Values:

- States like SP , RJ , and MG exhibit higher total order values, indicating higher sales volume in these regions.

•Varying Average Order Values:

- PB and RR Display higher average order values, indicating potentially higher-value purchases despite lower sales volume.

•Diverse Performance:

- Some states, such as AM , AC, and AP , have lower total values and moderate to higher average order values, suggesting varied market behaviours or lower purchase frequency.

•Significant Differences:

- SP stands out with the highest total order value but relatively lower average order value compared to other states, suggesting higher sales volume with potentially lower-priced items.

Recommendations:

Leverage High-Performing Markets: Focus on retaining and expanding markets with high total order values (SP, RJ, MG) to maintain market dominance and potentially increase sales volume.

Average Order Value Enhancement: Implement strategies to improve average order values, especially in states with lower averages, aiming to increase revenue per order.

Targeted Marketing: Develop targeted marketing campaigns or product offerings to attract higher-value orders in states displaying lower total but higher average values (PB, RR).

Customer Engagement: Emphasize customer engagement strategies to encourage repeat purchases and increase order frequency, particularly in states with lower total and average values.

##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.

```
SELECT
  order_id,
  TIMESTAMP_DIFF(order_delivered_customer_date,order_purchase_timestamp,DAY) AS `time_to_deliver`,
  TIMESTAMP_DIFF(order_estimated_delivery_date,order_delivered_customer_date,DAY) AS `diff_estimated_delivery`,
   CASE
       WHEN TIMESTAMP_DIFF(order_estimated_delivery_date,order_delivered_customer_date,DAY)>0 THEN "Good"
        WHEN TIMESTAMP_DIFF(order_estimated_delivery_date,order_delivered_customer_date,DAY)=0 THEN "Satisfied"
        WHEN TIMESTAMP_DIFF(order_estimated_delivery_date,order_delivered_customer_date,DAY)<0 THEN "Need to improve"
   END AS `delivery analysis`
FROM `target.orders`
WHERE LOWER(order_status)="delivered"
```

Row	order_id	time_to_deliver	diff_estimated_delive	delivery analysis
1	b60b53ad0bb7dacacf2989fe2...	12	-5	Need to improve
2	276e9ec344d3bf029ff83a161c...	43	-4	Need to improve
3	1a0b31f08d0d7e87935b819ed...	6	29	Good
4	cec8f5f7a13e5ab934a486ec9e...	20	40	Good
5	54e1a3c2b97fb0809da548a59...	40	-4	Need to improve
6	58527ee4726911bee84a0f42c...	10	48	Good
7	302bb8109d097a9fc6e9cefc5...	33	-5	Need to improve
8	10ed5499d1623638ee810eff1...	28	29	Good
9	cb837ba275cf8ffa9ded7e18f7...	12	-4	Need to improve
10	66057d37308e787052a32828...	38	-6	Need to improve
11	818996ea247803ddc123789f2...	9	35	Good

Insights:

Further analysis of the dataset revealed that more than 87,000 orders were delivered before the estimated delivery time, showcasing an optimized delivery process. Conversely, a small number of orders were slightly delayed beyond the estimated time. This analysis is solely based on the time difference between actual and estimated delivery times, without setting specific threshold values for these timeframes.

Recommendations:

Based on the insights derived from the analysis:

Recognize and Reward Efficiency: Acknowledge and incentivize the efficient delivery teams or processes that consistently deliver before the estimated time. This can boost morale and motivation among the teams.

Investigate Delayed Orders: Although few, the orders that were delivered slightly later than estimated warrant investigation. Understanding the reasons behind these delays, whether due to external factors or internal operational issues, can help implement corrective measures.

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

```
SELECT
  c.customer_state,
  ROUND(SUM(oi.freight_value)/COUNT(DISTINCT oi.order_id),2) AS `Top_5_avg_per_state`
 FROM `target.orders` o
JOIN `target.order_items` oi
ON o.order_id=oi.order_id
JOIN `target.customers` c
ON o.customer_id=c.customer_id
GROUP BY c.customer_state
ORDER BY Top_5_avg_per_state DESC
LIMIT 5;
```

Row	customer_state	Top_5_avg_per_state
1	RR	48.59
2	PB	48.35
3	RO	46.22
4	AC	45.52
5	PI	43.04

#States with Lowest freight values

```
SELECT
  c.customer_state,
  ROUND(SUM(oi.freight_value)/COUNT(DISTINCT oi.order_id),2) AS `Lowest_5_avg_per_state`
FROM `target.orders` o
JOIN `target.order_items` oi
ON o.order_id=oi.order_id
JOIN `target.customers` c
ON o.customer_id=c.customer_id
GROUP BY c.customer_state
ORDER BY Lowest_5_avg_per_state
LIMIT 5;
```

Row	customer_state	Lowest_5_avg_per_s
1	SP	17.37
2	MG	23.46
3	PR	23.58
4	DF	23.82
5	RJ	23.95

Insights:

- SP has the lowest average freight cost among the states analyzed, while RR has the highest.
- States with lower average freight values may indicate better logistical efficiency or proximity to distribution centers, whereas higher values might suggest logistical challenges or greater distance from shipping centers.
- Understanding these variations in freight costs across states can aid in logistics planning, potentially optimizing shipping strategies to reduce costs in areas with higher freight values and maintaining efficiency in areas with lower values.

Recommendations:

Distribution Optimization: Establish or reinforce distribution centers in regions with lower freight costs for efficient shipping.
Route Efficiency: Invest in route optimization to reduce transit times and costs in areas with higher freight expenses.
Local Partnerships: Form alliances with local logistics providers in regions with high freight values for better shipping rates.
Regular Monitoring: Continuously analyse freight expenses to adapt logistics strategies promptly.
Alternative Carriers: Experiment with different carriers or services to discover cost-effective options.
Technology Implementation: Implement advanced logistics tech for efficient route planning and inventory management.

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

```
SELECT
   c.customer_state,
   ROUND(
            SUM(TIMESTAMP_DIFF(o.order_delivered_customer_date,o.order_purchase_timestamp,DAY))/
            COUNT(DISTINCT  o.order_id),2) AS `Highest_delivery_time`
FROM target.orders o
JOIN target.customers c ON o.customer_id = c.customer_id
WHERE LOWER(order_status)="delivered"
GROUP BY c.customer_state
ORDER BY 2 DESC
LIMIT 5;
```

Row	customer_state	Highest_delivery_tim
1	RR	28.98
2	AP	26.73
3	AM	25.99
4	AL	24.04
5	PA	23.32

#States with lowest delivery time

```
SELECT
    c.customer_state,
    ROUND(
    SUM(TIMESTAMP_DIFF(o.order_delivered_customer_date,o.order_purchase_timestamp,DAY))/
    COUNT(DISTINCT o.order_id),2 ) AS `Lowest_delivery_time`
FROM target.orders o
JOIN target.customers c ON o.customer_id = c.customer_id
WHERE LOWER(order_status)="delivered"
GROUP BY c.customer_state
ORDER BY 2
LIMIT 5;
```

Row	customer_state	Lowest_delivery_tim
1	SP	8.3
2	PR	11.53
3	MG	11.54
4	DF	12.51
5	SC	14.48

Insights:

- Freight costs and delivery times can have an interconnected relationship in logistics.
- Efficient logistics operations can help reduce freight costs by optimizing routes, consolidating shipments, and minimizing unnecessary delays, which can consequently lead to faster delivery times.

Recommendations:

We can achieve quick deliveries by Optimizing logistics operations, leveraging distribution centres etc. Optimising freight values will automatically affect the delivery times.

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

```
SELECT
   c.customer_state,
   ROUND(
      SUM(TIMESTAMP_DIFF(o.order_delivered_customer_date,o.order_purchase_timestamp,DAY))/
      COUNT(DISTINCT  o.order_id),2) AS `avg_delivery_time`,
  ROUND(
       SUM(TIMESTAMP_DIFF(o.order_estimated_delivery_date,o.order_purchase_timestamp,DAY))/
       COUNT(DISTINCT o.order_id),2) AS `avg_estimated_delivery_time`,
  ROUND(
        SUM(TIMESTAMP_DIFF(o.order_estimated_delivery_date,o.order_purchase_timestamp,DAY))/
        COUNT(DISTINCT o.order_id)-
       SUM(TIMESTAMP_DIFF(o.order_delivered_customer_date,o.order_purchase_timestamp,DAY))/
        COUNT(DISTINCT o.order_id),2) AS `diff_time`
FROM `target.orders` o
JOIN `target.customers` c
ON c.customer_id=o.customer_id AND o.order_status="delivered"
GROUP BY c.customer_state
ORDER BY diff_time DESC
LIMIT 5;
```

Row	customer_state	avg_delivery_time	avg_estimated_deliv	diff_time
1	AC	20.64	40.73	20.09
2	RO	18.91	38.39	19.47
3	AP	26.73	45.87	19.13
4	AM	25.99	44.92	18.94
5	RR	28.98	45.63	16.66

Insights:

The presented table reveals a substantial disparity between the estimated delivery duration and the actual time taken for delivery, suggesting an efficient logistics system in place.

Recommendations:

Improve Estimated Delivery Estimations: The states generally experience significant gaps between estimated and actual delivery times. Refining estimated delivery projections could manage customer expectations better.

```
##6.Analysis based on the payments:
#6A.Find the month on month no. of orders placed using different payment types.
SELECT
    EXTRACT(YEAR FROM order_purchase_timestamp) AS `year`,
    -- FORMAT_DATETIME("%B", order_purchase_timestamp) AS `month`,
    EXTRACT(MONTH FROM order_purchase_timestamp) AS `month`,
    p.payment_type,
    COUNT(DISTINCT o.order_id) AS `no_of_orders`
FROM `target.orders` o
JOIN `target.payments` p
ON o.order_id=p.order_id
GROUP BY year,month,p.payment_type
ORDER BY year,month,p.payment_type
```

Row	year	month	payment_type	no_of_orders
1	2016	9	credit_card	3
2	2016	10	UPI	63
3	2016	10	credit_card	253
4	2016	10	debit_card	2
5	2016	10	voucher	11
6	2016	12	credit_card	1
7	2017	1	UPI	197
8	2017	1	credit_card	582
9	2017	1	debit_card	9
10	2017	1	voucher	33
11	2017	2	UPI	398
12	2017	2	credit_card	1347

Insights:

Trend in Payment Types: Over the months and years, credit card payments consistently dominate among the various payment types, showing a higher number of transactions compared to others.
Growth in Usage: There is a noticeable increase in transactions with time, indicating a growing customer base or increased usage of online transactions.
Seasonal Variations: Certain months exhibit fluctuations in the number of transactions across payment types, suggesting possible seasonal influences or specific promotional events that might affect consumer behavior.
Preferred Payment Modes: Credit card transactions maintain a steady lead across all periods, indicating a higher preference or ease of use compared to other payment methods like UPI, debit cards, or vouchers.
Diversity in Payment Modes: Although credit cards dominate, there is a diverse spread of payment methods being used, showcasing a multi-channel approach by customers for making payments.

Recommendations:

Promote Credit Card Offers: Given the dominance of credit card transactions, consider introducing exclusive offers, discounts, or reward programs specifically for credit card users to further incentivize this payment method.
Enhance UPI Experience: UPI transactions have shown consistent growth. Invest in improving the user experience for UPI payments by collaborating with payment service providers or offering cashback incentives to encourage its usage.
Target Seasonal Campaigns: Identify the months with high transaction volumes and design targeted marketing campaigns or promotions during these periods. Align offers with customer spending behaviors during these seasons.
Diversify Debit Card Usage: Though lesser in number compared to credit cards, promote the use of debit cards by offering benefits like cashback, discounts, or tie-ups with banks for exclusive deals.
Personalized Voucher Campaigns: Since vouchers are used, albeit in smaller numbers, tailor personalized voucher offers or discounts based on customer purchase history or preferences to encourage more voucher-based transactions.
Analysing Transaction Peaks: Investigate reasons behind spikes in certain months to capitalize on successful strategies and replicate them in other months for increased transactions across all payment types.

#6B.Find the no. of orders placed on the basis of the payment installments that have been paid.

```
SELECT
COUNT(DISTINCT order_id) AS `Total_EMI_Payments`
FROM `target.payments`
WHERE payment_installments>1;
```

Row	Total_EMI_Payments
1	51170

Insights:

Financial Stability: A high total EMI payment indicates a significant financial commitment and stability among customers or entities making these payments. It suggests a willingness and capability to manage and repay loans or financial obligations through EMIs.

Credit Usage: EMIs are commonly associated with credit purchases, indicating a preference for using credit facilities for larger purchases or investments rather than paying upfront.

Recommendations:

Financial Planning Services: Offer financial planning or advisory services to help individuals manage their EMI payments efficiently. This can include guidance on budgeting, debt management, or investment planning.

Loan Products: If you're a financial institution, consider promoting tailored loan products that align with customers' preferences for EMI payments. Highlight features such as flexible repayment options or lower interest rates to attract potential borrowers.

Customized Offerings: Analyze the data to understand patterns or preferences in EMI payments. Tailor promotions, offers, or financial products based on these insights to attract customers seeking EMI-based payment solutions.

Digital Payment Platforms: Introduce or enhance digital platforms for managing EMI payments, providing convenience, and encouraging timely repayments.