# Target SQL Case Study

**1. Checking data type of all columns in the “customers” table.**

Query

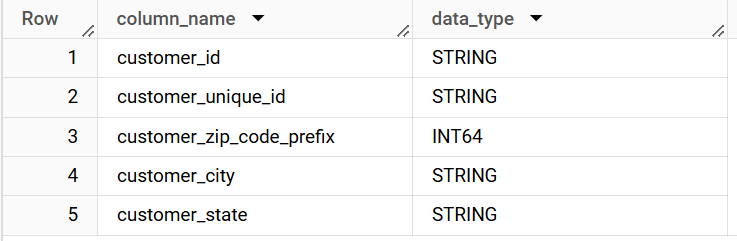
select column\_name,

data\_type

from target.INFORMATION\_SCHEMA.COLUMNS

where table\_name = 'customers'

Screenshot



Insights

Customers table contains 5 columns out of which the primary key(customer\_id) is of type STRING. All columns are of type STRING except customer\_zip\_code\_prefix which is of type INT

**2. Getting the time range between which the orders were placed.**

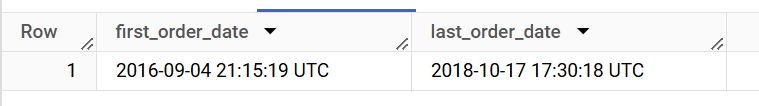
Query

select min(order\_purchase\_timestamp) as first\_order\_date,

max(order\_purchase\_timestamp) as last\_order\_date

from `target.orders`

Screenshot



Insights

This dataset contains orders in the date range 2016-09-04 to 2018-10-17.

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

Query

select g.geolocation\_city,

c.customer\_state,

count(\*) as Count\_of\_customers\_per\_city,

from target.customers c

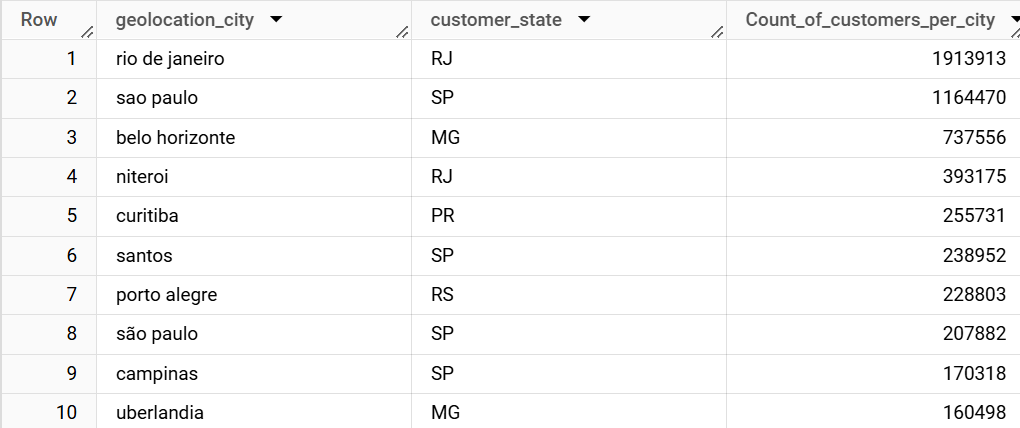
join target.geolocation g

on g.geolocation\_zip\_code\_prefix = c.customer\_zip\_code\_prefix

group by 1,2

order by 3 desc

Screenshot



Insights

The highest number of customers are coming from Rio De Janeiro and Sao Paulo. There are many cities which have only 1 customer.

**4. Checking if there is a growing trend in the no. of orders placed over the past years**

Query

WITH trend AS(

select EXTRACT(YEAR from order\_purchase\_timestamp) as  YEAR,

EXTRACT(MONTH from order\_purchase\_timestamp) as Month,

COUNT(\*) as order\_count

from `target.orders`

group by 1,2

order by 1,2

)

select \*,

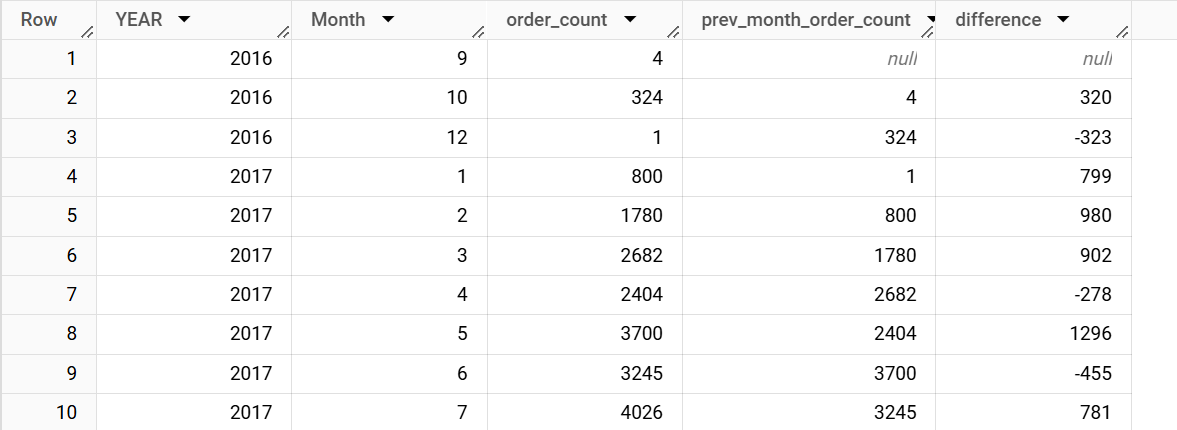
lag(order\_count,1) over(order by year,month) as prev\_month\_order\_count,

order\_count - lag(order\_count,1) over(order by year,month) as difference

from trend

order by year,month

Screenshot



Insight

As we can observe from the difference column the positive values indicate that there is a general increase in orders from the past months.

**5. Check if some kind of monthly seasonality is present in terms of the no. of orders being placed**

Query

select EXTRACT(MONTH from order\_purchase\_timestamp) as Month,

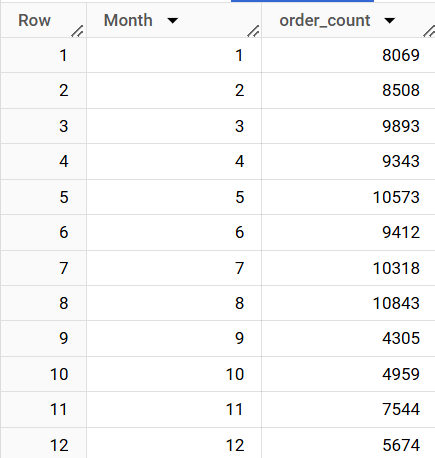
COUNT(\*) as order\_count

from `target.orders`

group by 1

order by 1

Screenshot



Insight

The number of orders peak during May, July and August. During September to December there is a dip in orders. There is a moderate number of orders during January to April. This could be due to holiday shopping. This suggests that there is a seasonality in the number of orders.

Action Items:

* Target can focus on special promotions during the September-December months to increase sales.

**6. 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

WITH time\_data AS

(

select \*,

CASE

  WHEN hour between 0 and 6 THEN 'Dawn'

  WHEN hour between 7 and 12 THEN 'Mornings'

  WHEN hour between 13 and 18 THEN 'Afternoon'

  ELSE 'Night'

END as time\_interval

from

(

select EXTRACT(HOUR from order\_purchase\_timestamp) as hour

from `target.orders`

)t

)

select time\_interval,

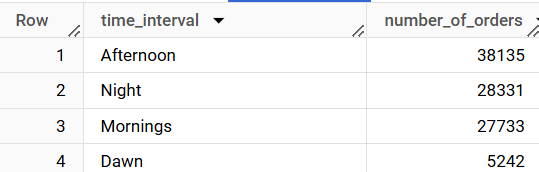
count(\*) as number\_of\_orders

from time\_data

group by time\_interval

order by 2 desc

Screenshot



Insights

Brazilians order the most during the Afternoon time with similar number of orders during the Night and Mornings and the least number of orders during Dawn.

Action Items

* Target can introduce sales or promotions during the Afternoon period to reach a wide customer base.
* They can increase the staff during the afternoon period to deal with the higher number of order.
* Target can allocate fewer resources during the Dawn time period to save on expenses.

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

Query

WITH inter AS

(

  select EXTRACT(YEAR from o.order\_purchase\_timestamp) as Year,

  EXTRACT(Month from o.order\_purchase\_timestamp) as Month, g.geolocation\_state,

  from`target.geolocation` g

  join `target.customers` c

  on g.geolocation\_zip\_code\_prefix = c.customer\_zip\_code\_prefix

  join `target.orders` o

  on o.customer\_id = c.customer\_id

)

select \*,

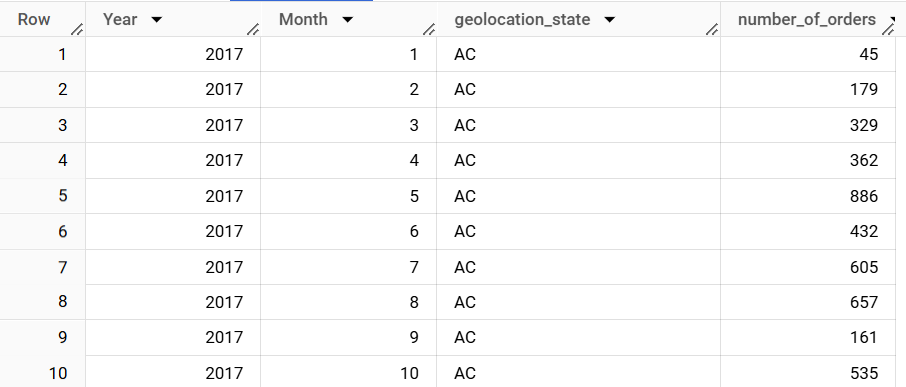
count(geolocation\_state) as number\_of\_orders

from inter

group by Year,Month,geolocation\_state

order by 3,1,2

Screenshot



Insights

Looking at the data for state AC. In 2017 there is a seasonality seen where in the summer months see peak number of orders. However, the order volume for state AC dipped in 2018 for all the months. We can find different observations for different states.

Action Items

* Target will need to investigate the reason for the dip in orders during 2018 for states like AC, to understand the reason for decreasing sales in this state.
* Target can check state specific patterns to localize promotions.
* If many states follow similar trends Target can provide broader promotions applying to many states.
* Target can increase staffing in states with high number of orders

**8. How are the customers distributed across all the states?**

Query

select customer\_state,

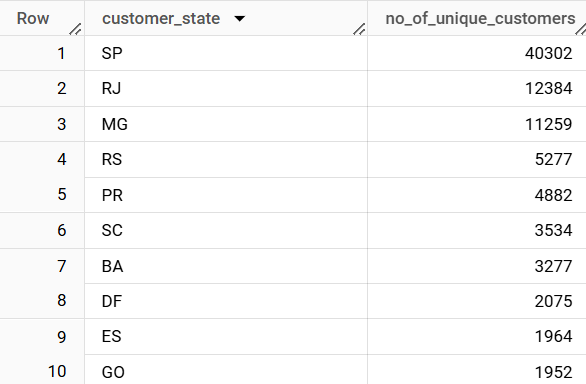
count(DISTINCT customer\_unique\_id) as no\_of\_unique\_customers

from `target.customers`

group  by customer\_state

order by 2 desc

Screenshot



Insights

We can observe that there is a huge number of unique customers in the states SP, RJ and MG. The least number of customers can be observed in the state RR, AP and AC with RR having the least.

Action Items

* Target will need to identify the reasons for low number of customers in RR,AP and AC. These can be limited accessibility to store which can be addressed. They will need to raise brand awareness in these areas.
* In the high customer states, Target can launch a loyalty programme to retain the customers and also introduce special discounts. They should also increase staffing in these states to ensure all customers are served on time.

**9. Find the % increase in the cost of orders from year 2017 to 2018**

Query

WITH yearly\_payments AS(

select EXTRACT(YEAR from 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 order\_purchase\_timestamp) in (2017,2018)

AND EXTRACT(MONTH from order\_purchase\_timestamp) in (1,2,3,4,5,6,7,8)

group by EXTRACT(YEAR from order\_purchase\_timestamp)

)

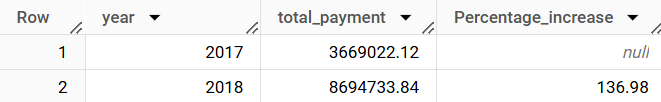
select year,total\_payment,

ROUND((total\_payment - LAG(total\_payment) over(order by YEAR))/LAG(total\_payment) over(order by YEAR) \*100,2)  as Percentage\_increase

from yearly\_payments

order by 1

Screenshot



Insights

The total payments made by customers has doubled in 2018 compared to 2017. There is a 137% increase in payment.

Action Items

* Target can introduce loyalty programmes to retain the new customers gained.

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

Query

select c.customer\_state,

ROUND(SUM(price),2) as Total\_price,

ROUND(AVG(price),2) as Avg\_price

from `target.order\_items` i

join `target.orders` o

on i.order\_id = o.order\_id

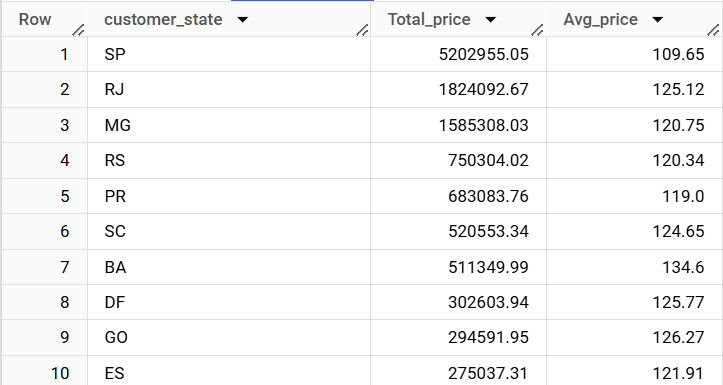
join `target.customers` c

on o.customer\_id = c.customer\_id

group by c.customer\_state

order by 2 desc,3 desc

Screenshot



Insights

The state with the highest total price is Sao Paulo (SP). We can see that it has a low average price as well which might contribute to more sales. The states with the lowest total price have high average price. This might explain the reason behind low number of customers here.

Action Items

* In states with high average prices and low total sales, the customers may be deterred by the pricing. Target may need to reassess their pricing strategies in these states.
* In states with high total sales and low average prices, target should try to sustain these prices as the customers are responding well to this.

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

Query

select c.customer\_state,

ROUND(SUM(i.freight\_value),2) as Total\_Freight\_value,

ROUND(AVG(i.freight\_value),2) as Avg\_Freight\_value

from `target.order\_items` i

join `target.orders` o

on i.order\_id = o.order\_id

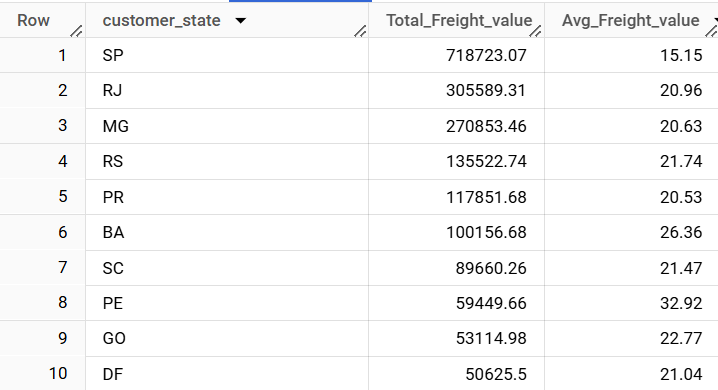
join `target.customers` c

on o.customer\_id = c.customer\_id

group by c.customer\_state

order by 2 desc,3 desc

Screenshot



Insights

The state with highest total freight value paid by customers also has the lowest average freight value. This can indicate a high volume of orders with low shipping costs. Conversely the state with lowest freight value has the highest average freight value cost. This indicates a low volume of orders with high shipping costs.

Action Items

* Target can try to optimize the costs in the low total volume states by giving discounts or a membership plan to minimize shipping costs for customers.

**12. Find the no. of days taken to deliver each order from the order’s purchase date as delivery time. Calculate the difference (in days) between the estimated & actual delivery date of an order.**

Query

select order\_id, date\_diff(order\_delivered\_customer\_date

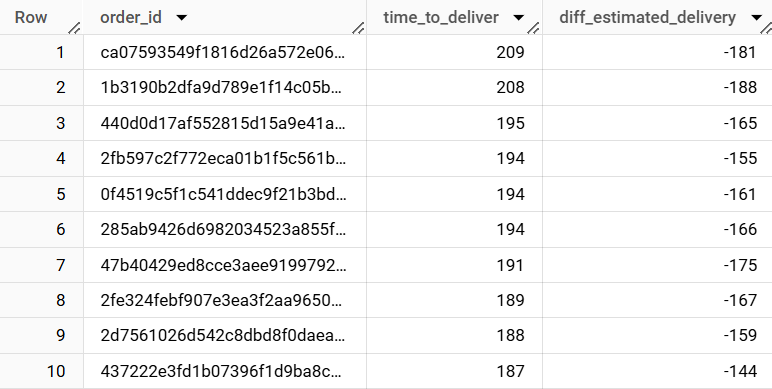
,order\_purchase\_timestamp,DAY) as time\_to\_deliver,

date\_diff(order\_estimated\_delivery\_date,order\_delivered\_customer\_date,DAY) as diff\_estimated\_delivery

from `target.orders`

order by 2 desc,3 desc

Screenshot



Insights

We can observe that there 6.5% of the cases where the order has been delayed i.e the order was delivered after the estimated arrival date. The average time to deliver was 12.09 days.

Action Items

* Target can improve their logistics to reduce delivery time to customer. They can set alerts if any delivery takes longer than the estimated date of arrival.

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

Query

(

select c.customer\_state,

ROUND(AVG(i.freight\_value),2) as Average\_Freight\_Value

from `target.order\_items` i

join `target.orders` o

on i.order\_id = o.order\_id

join `target.customers` c

on o.customer\_id = c.customer\_id

group by c.customer\_state

order by 2 asc

limit 5

)

UNION ALL

(

select c.customer\_state,

ROUND(AVG(i.freight\_value),2) as Average\_Freight\_Value

from `target.order\_items` i

join `target.orders` o

on i.order\_id = o.order\_id

join `target.customers` c

on o.customer\_id = c.customer\_id

group by c.customer\_state

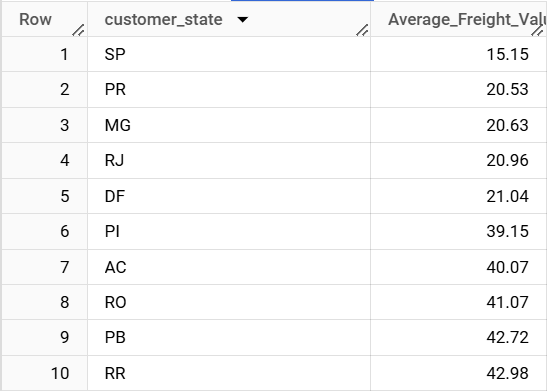
order by 2 desc

limit 5

)

order by Average\_Freight\_Value asc

Screenshot



Insights

States such as SP, PR, MG have low freight costs while RR, PB and RO have high freight costs. This could be due to the geographic position resulting in higher transportation costs.

Action Items

* Target can focus their marketing and promotion in lower cost states to increase their sales.
* They can try to order in bulk to reduce freight costs.

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

Query

WITH state\_delivery AS(

select c.customer\_state,

ROUND(AVG(date\_diff(order\_delivered\_customer\_date,order\_purchase\_timestamp,DAY)),2) as delivery\_time

from `target.orders` o

join `target.customers`c

on o.customer\_id = c.customer\_id

group by c.customer\_state

)

select \*

from

(

  (select \*

  from state\_delivery

  order by delivery\_time asc

  LIMIT 5)

  UNION ALL

  (select \*

  from state\_delivery

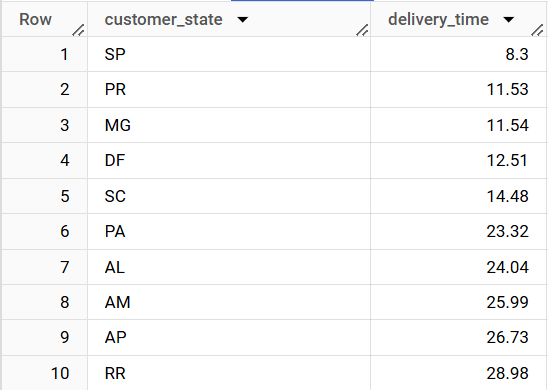
  order by delivery\_time desc

  LIMIT 5)

)

order by delivery\_time asc

Screenshot



Insights

States such as SP, PR and MG have relatively fast delivery times compared to RR, AP and AM. This could be due their geographic locations as we can see it is similar to the freight costs.

Action Items

* Target should try to improve the logistics for states with slow delivery.

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

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.orders` o

join `target.customers`c

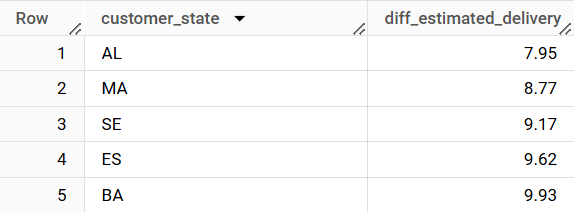
on o.customer\_id = c.customer\_id

where order\_delivered\_customer\_date IS NOT NULL

group by c.customer\_state

order by 2 asc

Screenshot



Insights

Action Items

**16. Find the month-on-month no. of orders placed using different payment types.**

Query

select EXTRACT(YEAR from o.order\_purchase\_timestamp) AS year,

EXTRACT(MONTH from o.order\_purchase\_timestamp) AS month,

count(CASE When lower(p.payment\_type) = 'credit\_card' Then 1 END) AS credit\_card\_orders,

count(CASE When lower(p.payment\_type) = 'debit\_card' Then 1 END) AS debit\_card\_orders,

count(CASE When lower(p.payment\_type) = 'voucher' Then 1 END) AS voucher\_orders,

count(CASE When lower(p.payment\_type) = 'upi' Then 1 END) AS upi\_orders,

count(CASE When lower(p.payment\_type) = 'not\_defined' Then 1 END) AS not\_defined\_orders

from `target.payments` p

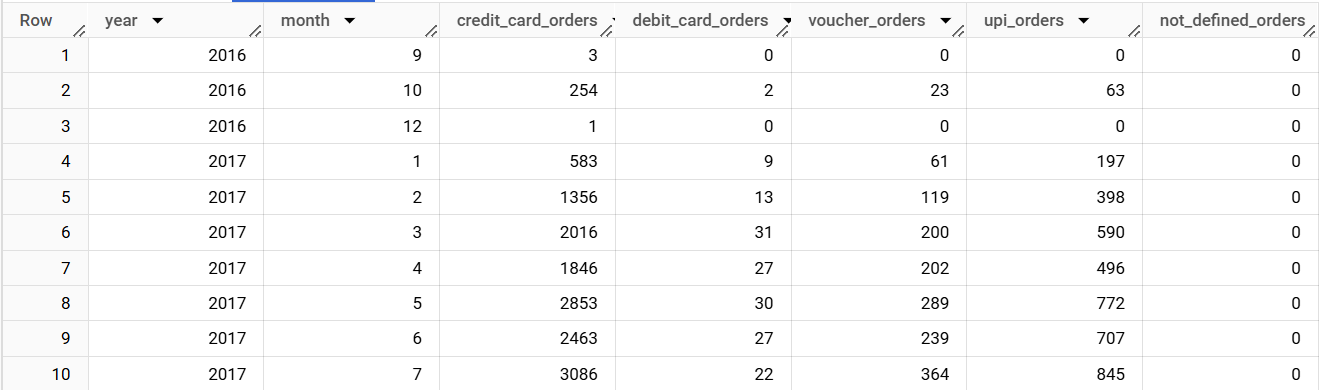
join `target.orders` o

on p.order\_id = o.order\_id

group by EXTRACT(YEAR from o.order\_purchase\_timestamp),EXTRACT(MONTH from o.order\_purchase\_timestamp)

order by 1,2

Screenshot



Insights

We can observe a steady increase in usage credit card and UPI as mode of payments. Debit card and voucher method of payments are low in number but are constant.

Action Items

* By observing the increasing trend in usage of credit cards, Target can partner with different banks to bring offers to increase sales.
* There is a growing trend in UPI payments. Target can ensure that all stores are equipped with QR codes for scanning and ease of use for customers to take advantage of this.
* The voucher payment method is stable thus Target can introduce loyalty programs for customers paying with vouchers.

**17. 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 number\_of\_orders

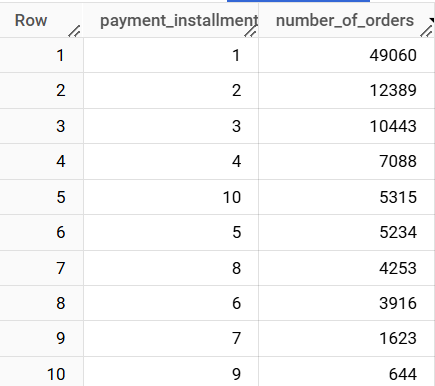
from `target.payments`

where payment\_installments>0

group by payment\_installments

order by 2 desc

Screenshot



Insights

Customers prefer paying the full amount in the least amount of installments possible. There is a sharp decline in paying for orders in more than 10 installments.

Action Items

* Target can focus on offers surrounding 1 or 2 installment payments as these are the most preferred by the customers.
* Try to reduce interest paid on longer installments to promote longer EMI payments.