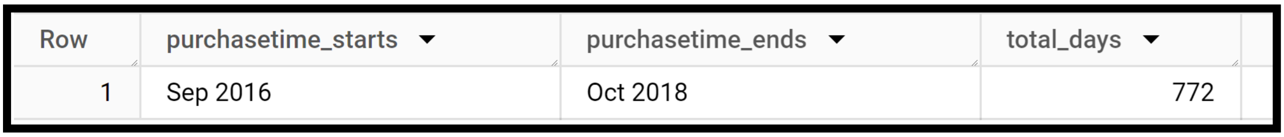
Target SQL Business Case

Submitted by: Krina Patel

Get the time range between which the orders were placed.

```
Select format_timestamp("%b %y",min(order_purchase_timestamp)) as purchasetime_starts,
          format_timestamp("%b %y",max(order_purchase_timestamp)) as purchasetime_ends,
          timestamp_diff(max(order_purchase_timestamp), min(order_purchase_timestamp), day)
as total_days
from `target.orders`
```

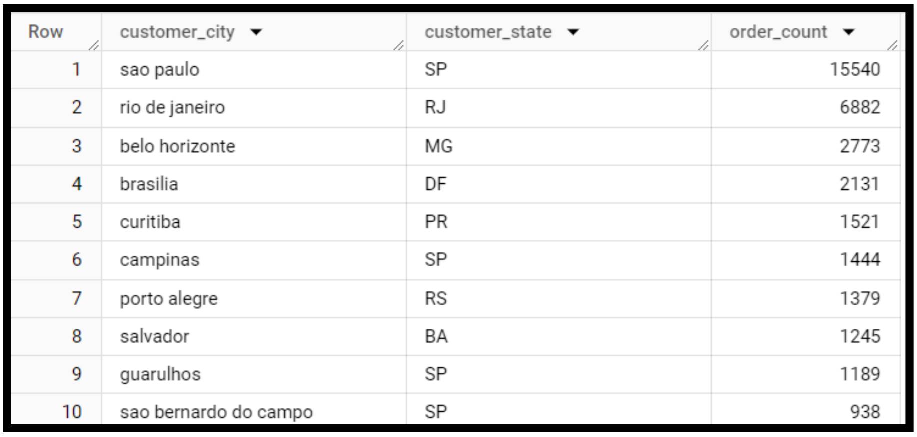


Row	purchasetime_starts	purchasetime_ends	total_days
1	Sep 2016	Oct 2018	772

Inference: Orders were placed between September 2016 and October 2018. Data of 772 days is available.

Count the Cities and states of customers who ordered during the given period.

```
Select distinct c.customer_city,
                      c.customer_state,
                     count(o.order_id) as order_count
from   `target.orders` o
join     `target.customers` c
on       o.customer_id = c.customer_id
group by c.customer_city, c.customer_state
order by order_count  desc
```



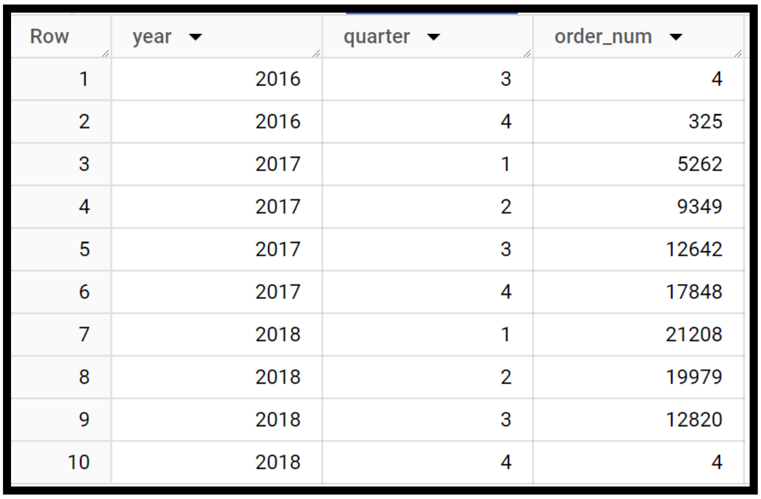
Row	customer_city	customer_state	order_count
1	sao paulo	SP	15540
2	rio de janeiro	RJ	6882
3	belo horizonte	MG	2773
4	brasilia	DF	2131
5	curitiba	PR	1521
6	campinas	SP	1444
7	porto alegre	RS	1379
8	salvador	BA	1245
9	guarulhos	SP	1189
10	sao bernardo do campo	SP	938

Inference: Sao Paulo and Rio de Janeiro led online retail sales in Brazil from 2016 to 2018.

Conclusion: The southern region is the richest in Brazil with a high population. While the Northeast region exhibits growth potential, representing about 13% of Brazil's e-commerce market, it faces challenges in terms of lower per capita income compared to the southern states. (1)

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

```
Select extract(year from order_purchase_timestamp) as year,

       extract(quarter from order_purchase_timestamp) as quarter,

       count(distinct order_id) as order_num

from target.orders

group by 1,2

order by 1,2
```



Row	year	quarter	order_num
1	2016	3	4
2	2016	4	325
3	2017	1	5262
4	2017	2	9349
5	2017	3	12642
6	2017	4	17848
7	2018	1	21208
8	2018	2	19979
9	2018	3	12820
10	2018	4	4

Inference: Monthly orders typically remain low during business setup and winding-down periods, so quarterly data is used for accuracy.

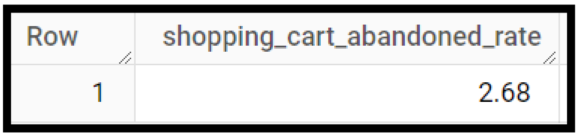
Conclusion: Order numbers showed consistent growth until the first quarter of 2018, with slight fluctuations in the second quarter. Growth was evident in every quarter from 2016 to 2017. However, assessing business growth solely based on profits may overlook non-monetary factors. So, the customer cart abandonment rate is estimated; however, it is approximate and not precise due to data limitations.

Approx. shopping cart abandonment rate

Shopping cart abandonment rate = (no. of unfinalized purchases/no. of total orders) * 100

Unfulfilled orders comprise those not delivered or 'processing' orders. This includes items that were created but not delivered, shipped but not received (and assumed canceled post-shipment), canceled but not delivered, and invoiced but not delivered (interpreted as abandoned shopping cart items). Processing orders are excluded as they are nearing delivery. Additionally, orders marked as unavailable and undelivered are considered unfinalized.

```
select ROUND((a.unfinalized_orders/b.total_order* 100),2) as shopping_cart_abandoned_rate
from
(select count(order_id) as unfinalized_orders
from `target.orders` o
where order_delivered_customer_date is null and order_status != 'processing')a,
(select count(order_id) as total_order from `target.orders`)b
```



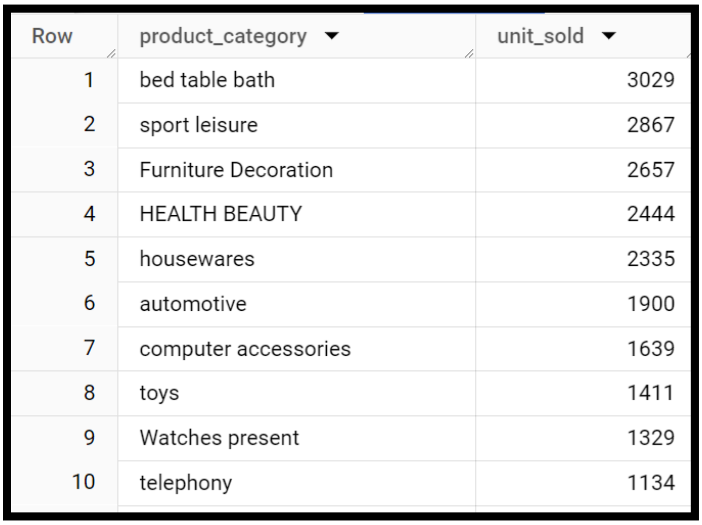
Row	shopping_cart_abandoned_rate
1	2.68

Inference: The approximate shopping cart abandonment rate is 2.68%, comparatively low, but lacks precise data.

Conclusion: In Brazil, around 15% of customers frequently abandon carts, with 72% due to high shipping costs (2). Further analysis is required to understand the shipping cost impact on cart abandonment and business growth.

No. of products sold in product category.

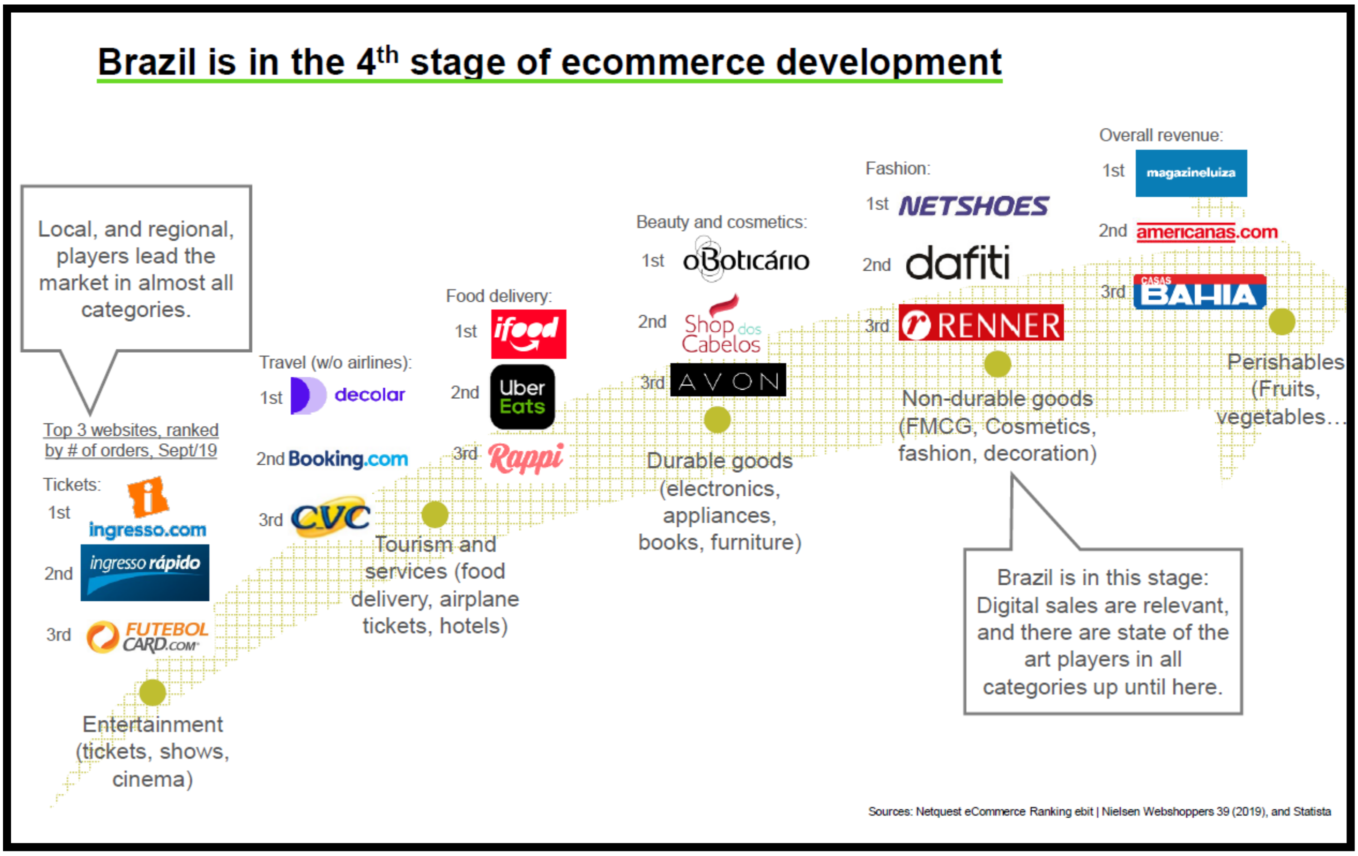
```
select product_category,
        count(product_id) as unit_sold
from `target.products`
where product_category is not null
group by product_category
order by unit_sold desc
```



Row	product_category	unit_sold
1	bed table bath	3029
2	sport leisure	2867
3	Furniture Decoration	2657
4	HEALTH BEAUTY	2444
5	housewares	2335
6	automotive	1900
7	computer accessories	1639
8	toys	1411
9	Watches present	1329
10	telephony	1134

Inference: Here top categories were home and décor, cosmetics, and sports leisure, which indicates that Brazilians purchase non-durable goods online and that places them in the 4th stage of e-commerce development (NetquesteCommerce Ranking ebit| Nielsen Webshoppers39 (2019), and Statista)

Conclusion: Based on the top-selling product, one can plan the inventory of those products and they may help in driving growth



Brazil is in the 4th stage of ecommerce development

Local, and regional, players lead the market in almost all categories.

Top 3 websites, ranked by # of orders, Sept/19

Tickets:
1st ingresso.com
2nd ingresso rápido
3rd FUTEBOL CARD.COM

Entertainment (tickets, shows, cinema)

Travel (w/o airlines):
1st decolar
2nd Booking.com
3rd CVC

Food delivery:
1st ifood
2nd Uber Eats
3rd Rappi

Tourism and services (food delivery, airplane tickets, hotels)

Beauty and cosmetics:
1st oBoticário
2nd Shop dos Cabelos
3rd AVON

Durable goods (electronics, appliances, books, furniture)

Fashion:
1st NETSHOES
2nd dafiti
3rd RENNER

Non-durable goods (FMCG, Cosmetics, fashion, decoration)

Overall revenue:
1st magazineluiza
2nd americanas.com
3rd CASAS BAHIA

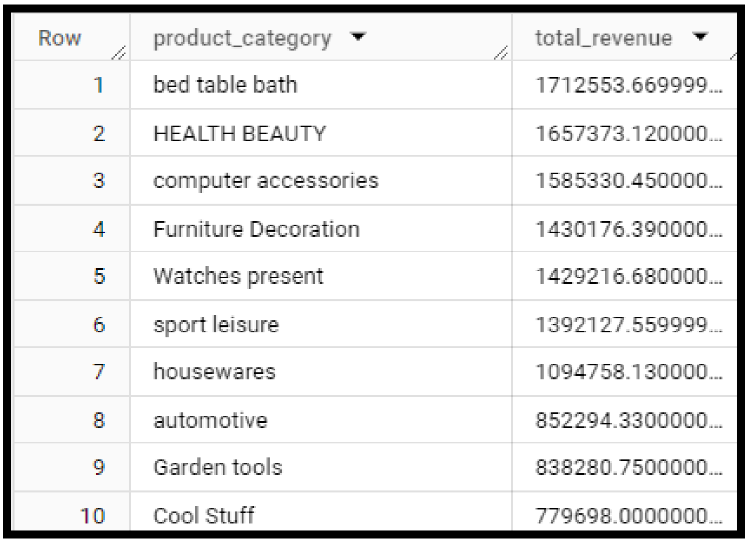
Perishables (Fruits, vegetables…

Brazil is in this stage: Digital sales are relevant, and there are state of the art players in all categories up until here.

Sources: Netquest eCommerce Ranking ebit | Nielsen Webshoppers 39 (2019), and Statista

Growth of product category based on revenue

```
select p.product_category,

       sum(pa.payment_value) as total_revenue

from `target.payments` pa

join `target.order_items` o

on

pa.order_id = o.order_id

join `target.products` p

on

o.product_id = p.product_id

group by p.product_category

order by total_revenue desc
```



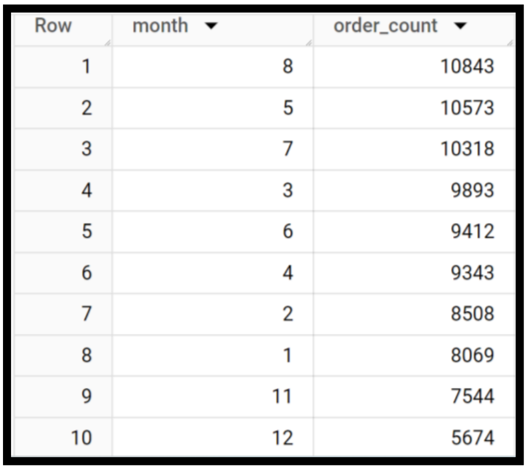
Row	product_category	total_revenue
1	bed table bath	1712553.669999…
2	HEALTH BEAUTY	1657373.120000…
3	computer accessories	1585330.450000…
4	Furniture Decoration	1430176.390000…
5	Watches present	1429216.680000…
6	sport leisure	1392127.559999…
7	housewares	1094758.130000…
8	automotive	852294.3300000…
9	Garden tools	838280.7500000…
10	Cool Stuff	779698.0000000…

Inference: Health and beauty ranks 4th by units sold but 2nd by total revenue.

Conclusion: Assessing product categories by revenue helps in identifying the most profitable products for planning purposes. While the factors mentioned indicate positive signs of business growth over the given period, metrics such as customer retention rate, customer lifetime value, customer churn rate, and conversion rate are necessary to confirm growth definitively.

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,

        count(order_id) as order_count

from

    `target.orders`

group by

month

order by order_count desc, month
```



Row	month	order_count
1	8	10843
2	5	10573
3	7	10318
4	3	9893
5	6	9412
6	4	9343
7	2	8508
8	1	8069
9	11	7544
10	12	5674

Inference: High order volumes occurred in the 8th, 5th, and 7th months and lower-order volumes in the 11th and 12th months.

Conclusion: Seasonal patterns influence order numbers

July and August: Strong sales due to Father's Day and mid-month vacations. (htt3)

May and June: High sales for Mother's Day and Lovers' Day.

March: Promotions on International Women's Day and Consumer Day drive sales.

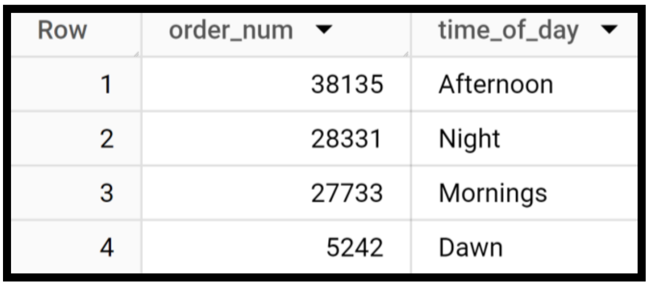
November-December: Typically, high sales for Cyber Monday and Black Friday; however, here the sales are low, which can be due to unknown micro or macro factors.

Macro factors can have an impact on the seasonality. For example, Black Friday is a big day sale but on the last Friday of November; in 2022, about 6.5 million checkouts were made on Black Friday in Brazil, which was fewer than the year before. This drop was because the Brazilian national team played their first World Cup game in Qatar on Thursday, November 24, which affected sales. (4)

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 count(order_id) as order_num,

    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'

        when extract(hour from order_purchase_timestamp) between 19 and 23 then 'night'

  end as time_of_day

from `target.orders` o

group by time_of_day

order by order_num desc
```



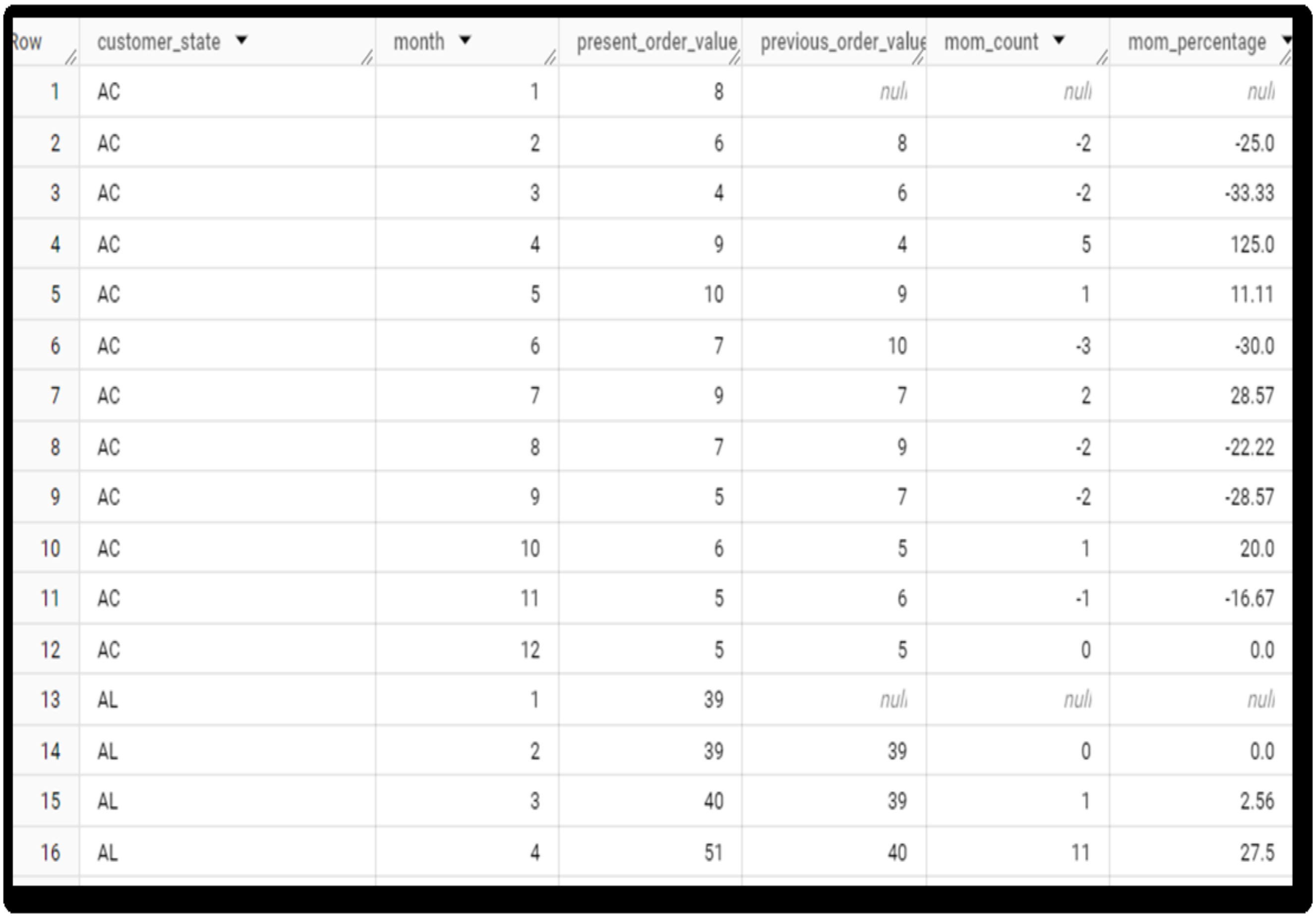
Row	order_num ▾	time_of_day ▾
1	38135	Afternoon
2	28331	Night
3	27733	Mornings
4	5242	Dawn

Inference: The highest number of orders are placed during the afternoon and night.

conclusion: Given the peak traffic during these times, effective order management is essential. To attract more customers during low-order time slots, consider utilizing email marketing with countdown timers. For instance, offering deals with a countdown timer indicating "offer ends in 4 hours”. it can generate interest, create urgency, and trigger quick responses due to the fear of missing out (FOMO). A case study found that integrating countdown timers into emails during the Black Friday sale resulted in a 400% higher conversion rate . (5)

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

```
 with cte as (select c.customer_state,
                            extract(month from order_purchase_timestamp) as month,
                            count(order_id) as present_order_value
                            from `target.orders` o
                             join `target.customers` c
                             on
                            o.customer_id = c.customer_id
                            group by 1,2)


select customer_state,
          month,
          present_order_value,
          lag(present_order_value) over(partition by customer_state order by month) as
previous_order_value,
          present_order_value - (lag(present_order_value) over(partition by customer_state order
by month)) as mom_count,
          round(((present_order_value - lag(present_order_value) over(partition by
customer_state order by month)) / lag(present_order_value) over(partition by customer_state
order by month))* 100,2) as mom_percentage
from cte
order by 1,2
```



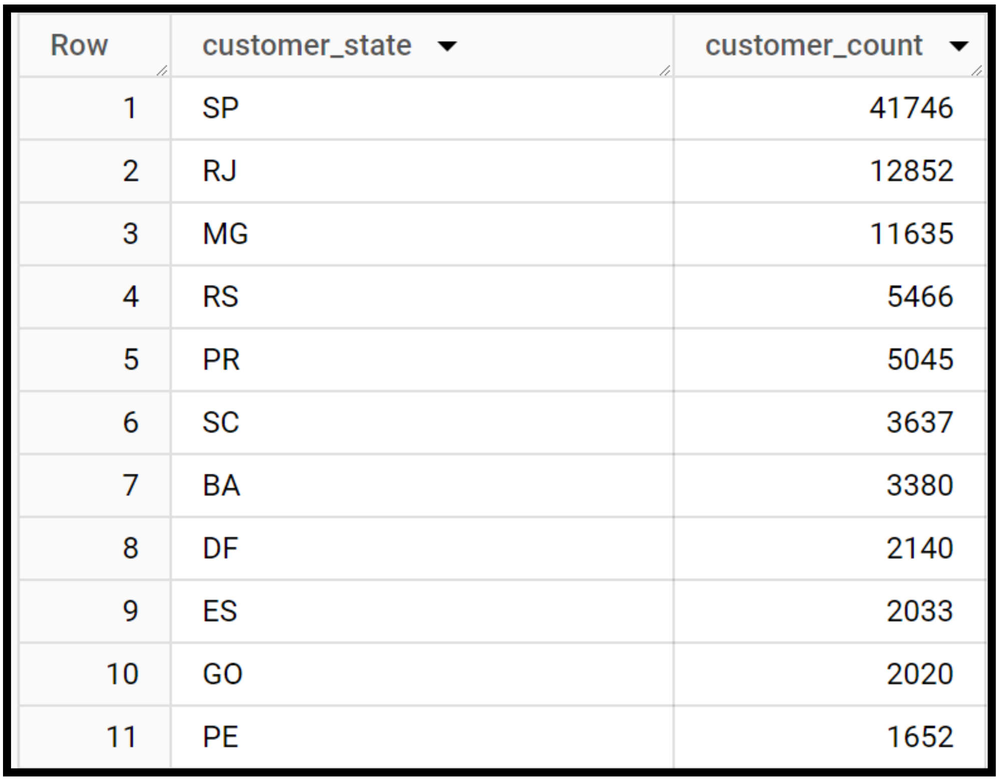
Row	customer_state	month	present_order_value	previous_order_value	mom_count	mom_percentage
1	AC	1	8	null	null	null
2	AC	2	6	8	-2	-25.0
3	AC	3	4	6	-2	-33.33
4	AC	4	9	4	5	125.0
5	AC	5	10	9	1	11.11
6	AC	6	7	10	-3	-30.0
7	AC	7	9	7	2	28.57
8	AC	8	7	9	-2	-22.22
9	AC	9	5	7	-2	-28.57
10	AC	10	6	5	1	20.0
11	AC	11	5	6	-1	-16.67
12	AC	12	5	5	0	0.0
13	AL	1	39	null	null	null
14	AL	2	39	39	0	0.0
15	AL	3	40	39	1	2.56
16	AL	4	51	40	11	27.5

Inference: Each state has month-wise some fluctuation in number of orders. No single state has a continuous growth trend.

Conclusion: During high-order month, implement customer retention strategies like personalized customer experiences, loyalty programs, and social media engagement. During low-demand months, provide offers and discounts, offer more delivery and return options, and run limited-period offers.

How are the customers distributed across all the states?

```
select customer_state,
        count(customer_unique_id) as customer_count
from `target.customers`
group by 1
order by 2 desc
```



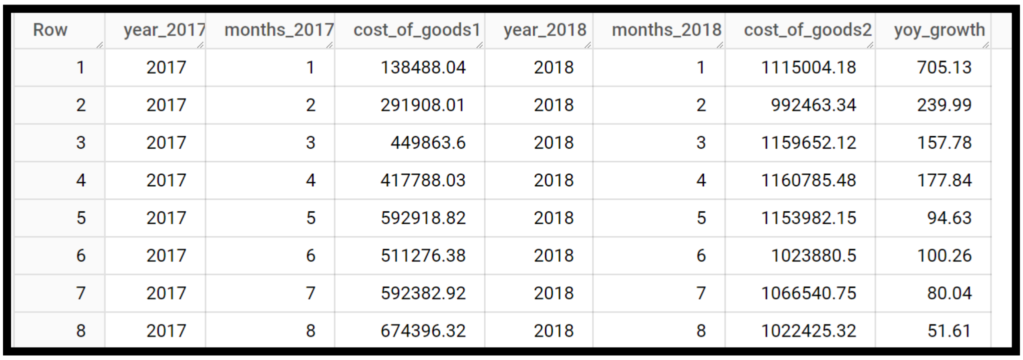
Row	customer_state	customer_count
1	SP	41746
2	RJ	12852
3	MG	11635
4	RS	5466
5	PR	5045
6	SC	3637
7	BA	3380
8	DF	2140
9	ES	2033
10	GO	2020
11	PE	1652

Inference: State SP has the highest number of customers followed by RJ and MG.

Conclusion: For the state with the lower number of orders, better logistic planning is required to ship products on time and at very low or no shipping cost at all. Geographic segmentation may help when combined with other customer attributes and psychographic segmentation as well

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

```
select *,
  round(((cost_of_goods2-cost_of_goods1)/cost_of_goods1)*100,2) as yoy_growth
from
(select
  extract(year from order_purchase_timestamp) as year_2017,
  extract(month from order_purchase_timestamp) as months_2017,
  round(sum(p.payment_value),2) as cost_of_goods1
from `target.orders` o

join `target.payments` p

on

o.order_id = p.order_id

where (extract(year from order_purchase_timestamp) = 2017) and extract(month from order_purchase_timestamp) between 1 and 8

group by 1,2)t1

join

(select extract(year from order_purchase_timestamp) as year_2018,

      extract(month from order_purchase_timestamp) as months_2018,

      round(sum(p.payment_value),2) as cost_of_goods2

 from `target.orders` o

 join `target.payments` p

 on

 o.order_id = p.order_id

where (extract(year from order_purchase_timestamp) = 2018) and extract(month from order_purchase_timestamp) between 1 and 8

group by 1,2)t2

on

t1.months_2017 = t2.months_2018

order by t1.months_2017
```



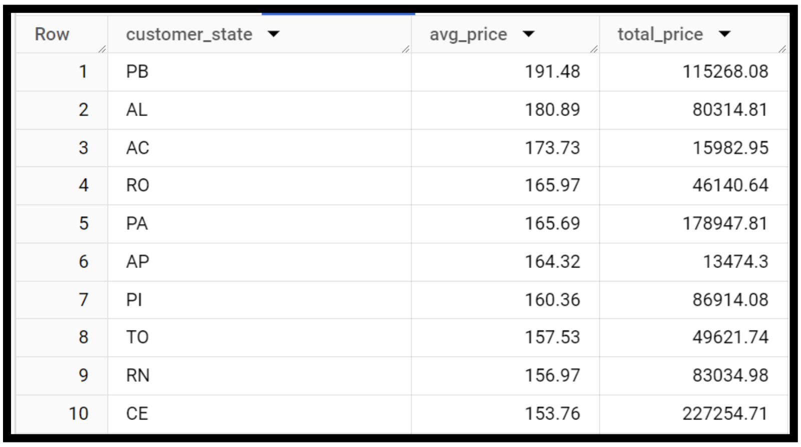
Row	year_2017	months_2017	cost_of_goods1	year_2018	months_2018	cost_of_goods2	yoy_growth
1	2017	1	138488.04	2018	1	1115004.18	705.13
2	2017	2	291908.01	2018	2	992463.34	239.99
3	2017	3	449863.6	2018	3	1159652.12	157.78
4	2017	4	417788.03	2018	4	1160785.48	177.84
5	2017	5	592918.82	2018	5	1153982.15	94.63
6	2017	6	511276.38	2018	6	1023880.5	100.26
7	2017	7	592382.92	2018	7	1066540.75	80.04
8	2017	8	674396.32	2018	8	1022425.32	51.61

Inference: The data shows growth for eight months, but it slows down by the end of the second quarter.

Conclusion: It's important to analyze this data for future marketing plans and to tailor marketing strategies locally. For example, Walmart didn't act locally and had to leave the Brazilian retail market.

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

```
select customer_state,

        avg(price) as avg_price,

        sum(price) as total_price

from `target.customers` c

join `target.orders` o

on c.customer_id = o.customer_id

join `target.order_items` oi

on

o.order_id = oi.order_id

group by customer_state

order by avg_price desc, total_price desc
```



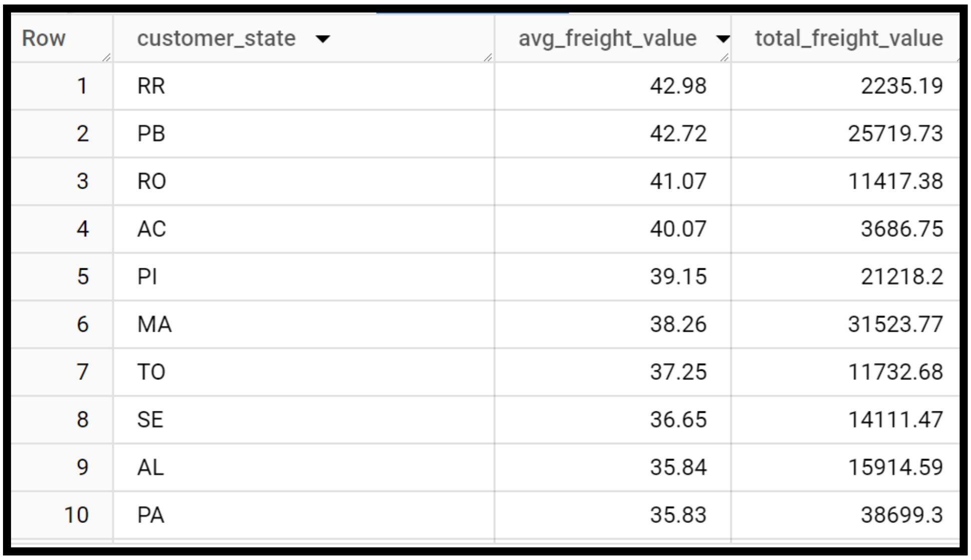
Row	customer_state	avg_price	total_price
1	PB	191.48	115268.08
2	AL	180.89	80314.81
3	AC	173.73	15982.95
4	RO	165.97	46140.64
5	PA	165.69	178947.81
6	AP	164.32	13474.3
7	PI	160.36	86914.08
8	TO	157.53	49621.74
9	RN	156.97	83034.98
10	CE	153.76	227254.71

Inference: Some states with the highest average values of order price aren't the same as those with the most orders, suggesting impact of the factors like higher revenue per customer, profitable product categories, or other factors such as high conversion rates or customer retention.

Conclusion: To improve average order values in states with lower averages, strategies like customer segmentation, cross-selling, upselling, offering low shipping costs or free delivery, and using automated product recommendations can be effective.

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

```
Select     customer_state,

           round(avg(freight_value),2) as avg_freight_value,
           round(sum(freight_value),2) as total_freight_value
           from `target.customers` c
           join `target.orders` o
            on c.customer_id = o.customer_id
            join `target.order_items` oi
           on
           o.order_id = oi.order_id
           group by customer_state
           order by avg_freight_value desc, total_freight_value desc
```



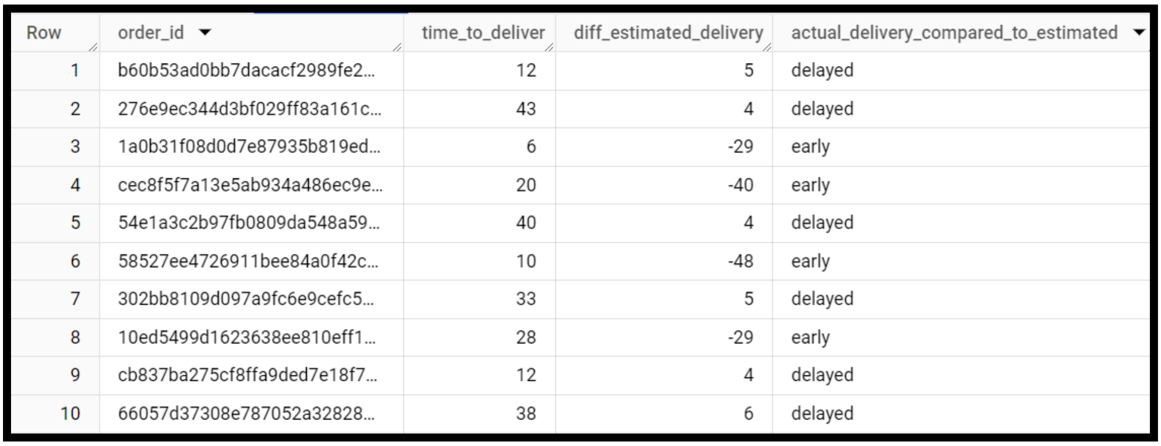
Row	customer_state	avg_freight_value	total_freight_value
1	RR	42.98	2235.19
2	PB	42.72	25719.73
3	RO	41.07	11417.38
4	AC	40.07	3686.75
5	PI	39.15	21218.2
6	MA	38.26	31523.77
7	TO	37.25	11732.68
8	SE	36.65	14111.47
9	AL	35.84	15914.59
10	PA	35.83	38699.3

Inference: Both the top five and bottom five freight values are significantly higher than the average freight value for the years 2017 to 2019, which ranged from 16.87 reals to 19.34 reals. (6)

Conclusion: High shipping costs lead to increased shopping cart abandonment rates and lower conversion rates. To address this, it's essential to either lower shipping costs or offer free delivery

Find the no. of days taken to deliver each order from the order's purchase date as delivery time.

```
with cte as (select  order_id,
                date_diff(order_delivered_customer_date,order_purchase_timestamp,day) as
time_to_deliver,
             date_diff(order_delivered_customer_date,order_estimated_delivery_date,day) as
diff_estimated_delivery
             from `target.orders`
            where order_delivered_customer_date is not null and order_status = 'delivered')
select      order_id,
            time_to_deliver,
            diff_estimated_delivery,
            case
                   when diff_estimated_delivery > 0 then 'delayed'
                   when diff_estimated_delivery < 0 then 'early'
                    when diff_estimated_delivery = 0 then 'same day'
                   end as actual_delivery_compared_to_estimated
             from cte
```



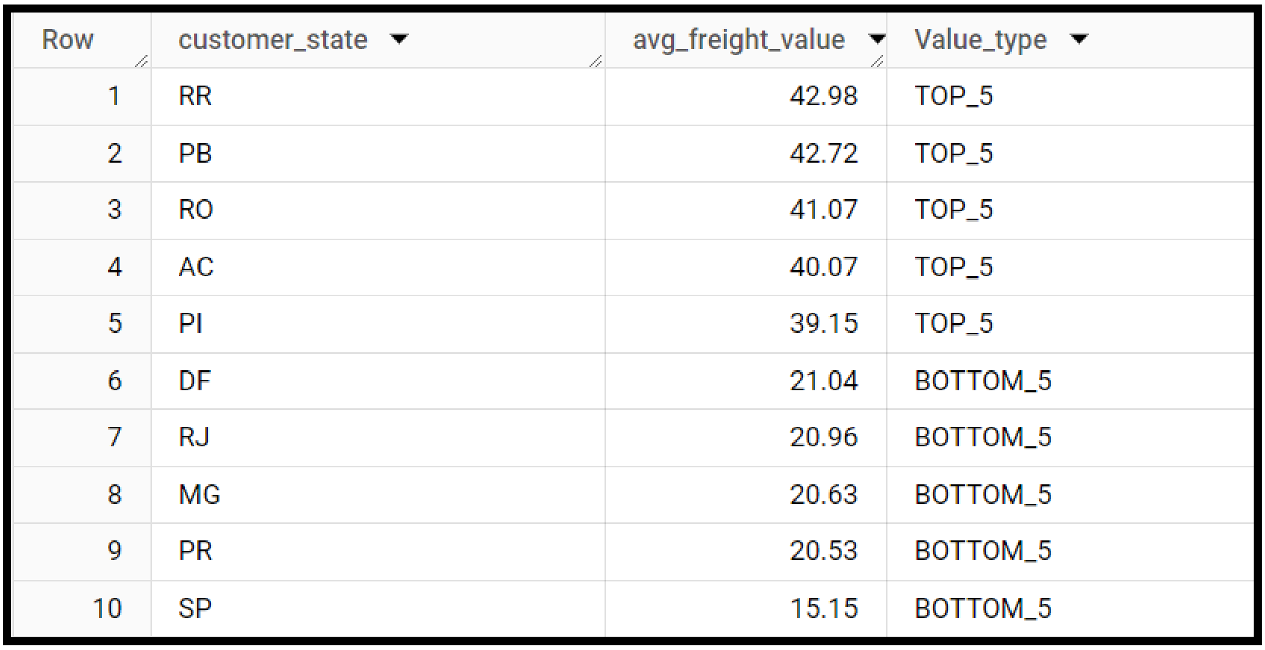
Row	order_id	time_to_deliver	diff_estimated_delivery	actual_delivery_compared_to_estimated
1	b60b53ad0bb7dacacf2989fe2...	12	5	delayed
2	276e9ec344d3bf029ff83a161c...	43	4	delayed
3	1a0b31f08d0d7e87935b819ed...	6	-29	early
4	cec8f5f7a13e5ab934a486ec9e...	20	-40	early
5	54e1a3c2b97fb0809da548a59...	40	4	delayed
6	58527ee4726911bee84a0f42c...	10	-48	early
7	302bb8109d097a9fc6e9cefc5...	33	5	delayed
8	10ed5499d1623638ee810eff1...	28	-29	early
9	cb837ba275cf8ffa9ded7e18f7...	12	4	delayed
10	66057d37308e787052a32828...	38	6	delayed

Inference: Some orders were delayed compared to the estimated delivery time, while others arrived early.

Conclusion: Delayed orders can negatively impact customer satisfaction levels and potentially lower the net promoter score, affecting future business growth and customer retention rates. It's crucial to view shipping as a service provided to customers rather than just passing the delivery cost to them. Proper inventory planning and shipping strategies can help mitigate the problem of delayed deliveries.

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

```
with cte as (select  c.customer_state,
                    round(avg(freight_value),2) as avg_freight_value,
                    dense_rank()over (order by round(avg(freight_value),2) desc) as
                    highest_freight_value,
                    dense_rank()over (order by round(avg(freight_value),2) asc) as
                    lowest_freight_value
            from `target.customers` c
            join `target.orders` o
            on c.customer_id = o.customer_id
            join `target.order_items` oi
            on
            o.order_id = oi.order_id
            group by customer_state
              )
select customer_state,
      avg_freight_value,
      case
        when highest_freight_value <= 5 then 'TOP_5'
        when lowest_freight_value <= 5 then 'BOTTOM_5'
      end as value_type
from cte
where highest_freight_value <= 5  or lowest_freight_value <= 5
order by avg_freight_value desc
```



Row	customer_state	avg_freight_value	Value_type
1	RR	42.98	TOP_5
2	PB	42.72	TOP_5
3	RO	41.07	TOP_5
4	AC	40.07	TOP_5
5	PI	39.15	TOP_5
6	DF	21.04	BOTTOM_5
7	RJ	20.96	BOTTOM_5
8	MG	20.63	BOTTOM_5
9	PR	20.53	BOTTOM_5
10	SP	15.15	BOTTOM_5

Inference: SP is a state with significantly high order numbers, and that's why it has the lowest freight value due to the economy of scale, even lower than the average e-commerce freight value from 2017 to 2019, which ranged approximately from 16.87 to 19.34 reals.

Conclusion: To reduce freight costs, consider these strategies:

1. Establish collection points midway between distant delivery points for easier customer access.
2. Use lightweight packaging to lower shipping costs influenced by item weight.
3. Offer reduced shipping costs for multiple-item purchases to decrease the number of deliveries.

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

```
with cte as (select  c.customer_state,

       ceil(avg(date_diff(order_delivered_customer_date,order_purchase_timestamp,day))) as
avg_delivery_time,

        dense_rank() over(order by
ceil(avg(date_diff(order_delivered_customer_date,order_purchase_timestamp,day))) desc) as
highest_delivery_time,

       dense_rank() over(order by
ceil(avg(date_diff(order_delivered_customer_date,order_purchase_timestamp,day))) asc) as
lowest_delivery_time

from `target.customers` c

join `target.orders` o
on c.customer_id = o.customer_id
join `target.order_items` oi
on
o.order_id = oi.order_id
group by customer_state
)
select customer_state,
       avg_delivery_time,
       case
          when highest_delivery_time  <= 5 then 'top_5'
          when lowest_delivery_time <= 5 then 'bottom_5'
       end as value_type
from cte
where highest_delivery_time <= 5  or lowest_delivery_time <= 5
order by avg_delivery_time desc, customer_state
```

Methodology: 1) Here average days values are converted in to the ceiling value, e.g days = 3.34 would be more than 3 days, anything above 3 means 4th day has already been started.

2) Instead of order by , dense_rank() is used , because it should be top 5 states based on the top 5 avg. delivery times.

Here top 5 avg. delivery times are 28,26,24,22,21 days and some of them repeat for more than one state and I wanted to consider all of them. So total 10 states will be qualified for top 5 states.

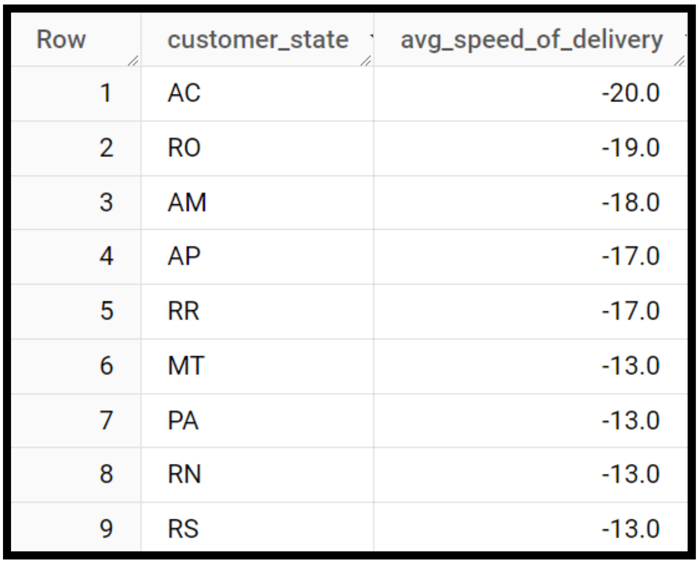
Inference: Average delivery time is quite high compared to customer's expectations and prevalent avg.delivery time in brazil.. In 2017 the average delivery time in Brazil was around 8 days. (7)

Conclusion: Brazilian buyers usually have to wait long periods to receive goods as the average delivery time is high, in this scenario delivering excellent shipping service is an opportunity to stand out in a growing and competitive market

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

```
with cte as (select  c.customer_state,


ceil(avg(date_diff(order_delivered_customer_date,order_estimated_delivery_date,day))) as
avg_speed_of_delivery,

                    dense_rank() over(order by
ceil(avg(date_diff(order_delivered_customer_date,order_estimated_delivery_date,day)))) as
d_rnk

            from `target.customers` c

            join `target.orders` o

           on c.customer_id = o.customer_id

           join `target.order_items` oi

           on

           o.order_id = oi.order_id
```

```
        where order_delivered_customer_date is not null and order_status = 'delivered'
        group by customer_state
        order by avg_speed_of_delivery,customer_state)
select   customer_state,
       avg_speed_of_delivery
       from cte
 where d_rnk <= 5
```



Row	customer_state	avg_speed_of_delivery
1	AC	-20.0
2	RO	-19.0
3	AM	-18.0
4	AP	-17.0
5	RR	-17.0
6	MT	-13.0
7	PA	-13.0
8	RN	-13.0
9	RS	-13.0

Methodology: 1) Here average days values are converted into the ceiling value, e.g days = 3.1234 would be more than 3 days, anything above 3 means 4th day has already been started.

2) Instead of order by, dense_rank() is used, because it should be top 5 states based on the top 5 avg. speed of delivery.

Here top 5 average speed of delivery times are -20, -19, -18, -17, -13 days and some of them repeat for more than one state and I wanted to consider all of them. So, total 9 states will be qualified for the top 5 states.

Inference : Here the negative value of the average speed of delivery shows that the product was delivered earlier than expected. E.g. earliest delivery is 20 days earlier than expected.

Conclusion: Earlier-than-expected delivery is one of the contributors to a high customer satisfaction rate, though it is not possible to maintain that in every case, so providing different shipping methods for different timeframes can be a win-win for both the customer and the business.

e.g. standard delivery costs 8 reals

premium 3 days business days costs 16 reals

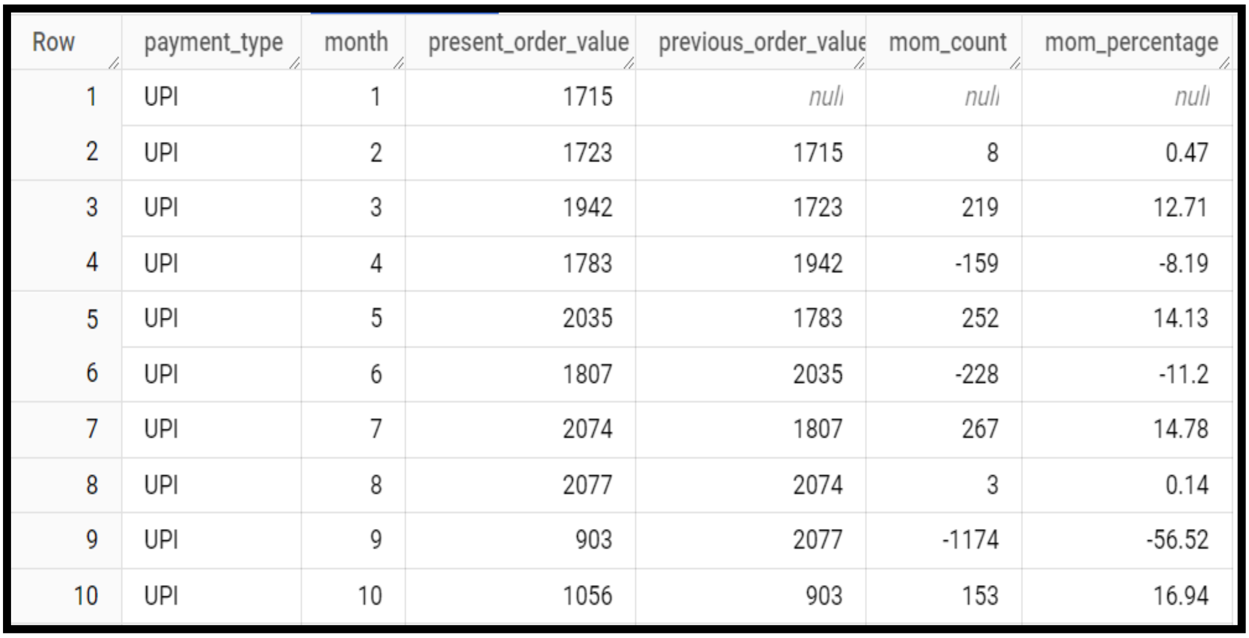
Express delivery 2 business days costs 29 reals

Expediated delivery 1 business day costs 42 reals

Note: costs here are hypothetical figures.

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

```
with cte as (select p.payment_type,
        extract(month from order_purchase_timestamp) as month,
        count(o.order_id) as present_order_value
from `target.orders` o
join `target.payments` p
on
o.order_id = p.order_id
group by 1,2)


select payment_type,
          month,
         present_order_value,
        lag(present_order_value) over(partition by payment_type order by month) as previous_order_value,
        present_order_value - (lag(present_order_value) over(partition by payment_type order by month)) as mom_count,
       round(((present_order_value - lag(present_order_value) over(partition by payment_type order by month)) / lag(present_order_value) over(partition by payment_type order by month))* 100,2) as mom_percentage
from cte
where payment_type != 'not_defined'
order by 1,2
```



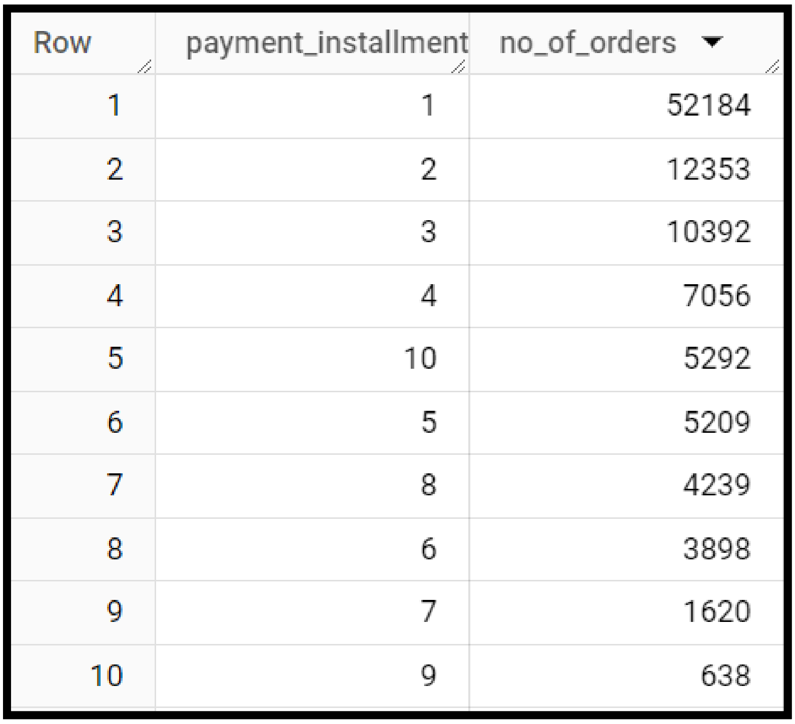
Row	payment_type	month	present_order_value	previous_order_value	mom_count	mom_percentage
1	UPI	1	1715	null	null	null
2	UPI	2	1723	1715	8	0.47
3	UPI	3	1942	1723	219	12.71
4	UPI	4	1783	1942	-159	-8.19
5	UPI	5	2035	1783	252	14.13
6	UPI	6	1807	2035	-228	-11.2
7	UPI	7	2074	1807	267	14.78
8	UPI	8	2077	2074	3	0.14
9	UPI	9	903	2077	-1174	-56.52
10	UPI	10	1056	903	153	16.94

Inference: Order value based on payment type keeps fluctuating over months.

Conclusion: Most of the customers in Brazil either use credit cards or UPI payment methods. Easy payment is one of the important aspects of a high conversion rate in Brazil.

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

```
Select  p.payment_installments,
        count(o.order_id) as no_of_orders
from `target.orders` o
join `target.payments` p
on
o.order_id = p.order_id
where order_status != 'canceled'
group by p.payment_installments
order by no_of_orders desc
```



Row	payment_installment	no_of_orders ▼
1	1	52184
2	2	12353
3	3	10392
4	4	7056
5	10	5292
6	5	5209
7	8	4239
8	6	3898
9	7	1620
10	9	638

Inference: Here approx. 49% of total orders are placed using more than one payment installment.

Conclusion: In Brazil, 36% of digital commerce users are from the low-income category and that's why most of all online retailers offer installment options. In 2020, approx. 27% of orders were in the category of 4-12 installments, and 21 % of orders were in the category of 2-3 installments. (9)

References:

(n.d.). Retrieved from (https://www.statista.com/statistics/770077/e-commerce-brazil-buyers-region/)

(n.d.). Retrieved from (https://www.statista.com/statistics/770077/e-commerce-brazil-buyers-region/)

(n.d.). Retrieved from https://www.statista.com/statistics/804172/brazil-net-promoter-score-marisa/

(n.d.). Retrieved from (https://www.ecommerce-nation.com/how-brazilians-shop-and-how-e-commerce-merchants-can-reach-them/)

(n.d.). Retrieved from https://www.salecycle.com/blog/strategies/countdown-timers-can-used-drive-ecommerce-sales/).

(n.d.). Retrieved from https://www.statista.com/statistics/779506/black-friday-e-commerce-sales-number-checkouts-brazil/

(n.d.). Retrieved from https://www.statista.com/statistics/783469/online-shopping-cart-abandonment-rate-reason-brazil/)

(n.d.). Retrieved from https://www.statista.com/statistics/769924/e-commerce-brazil-shipping-cost-checkout/

(n.d.). Retrieved from https://www.pagbrasil.com/insights/brazilian-e-commerce-trends-2017/

(n.d.). Retrieved from pymnts.com

(n.d.).

(n.d.). *NetquesteCommerce Ranking ebit| Nielsen Webshoppers39 (2019), and Statista.*