

## TARGET BRAZIL CASE STUDY

1(1) Data type of all columns in the "customers" table.

```
SELECT
```

```
    column_name,data_type
```

```
FROM target_brazil.INFORMATION_SCHEMA.COLUMNS
```

```
WHERE table_name='customers'
```

Query results		
JOB INFORMATION		RESULTS
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

Inference-After performing the above query I found the data type of different columns as- customer\_id,customer\_unique\_id , customer\_city,customer\_state has string data type where as customer\_zip\_code\_prefix has int as data type.

1(2) Get the time range between which the orders were placed.

```
SELECT DISTINCT
```

```
    MIN(order_purchase_timestamp) as start_time,
```

```
    MAX(order_purchase_timestamp) as end_time
```

```
FROM `target_brazil.orders`
```

Query results				
JOB INFORMATION		RESULTS	CHART	JSON
Row	start_time ▼	end_time ▼		
1	2016-09-04 21:15:19 UTC	2018-10-17 17:30:18 UTC		

Inference-after performing the above query for this condition I found that order were placed from 4 september 2016 to 17 october 2018.

1(3) Count the Cities & States of customers who ordered during the given period.

```
SELECT COUNT(DISTINCT customer_city) as city_count,
COUNT(DISTINCT customer_state) as state_count
FROM `target_brazil.customers` c
JOIN `target_brazil.orders` o
ON c.customer_id = o.customer_id
WHERE o.order_purchase_timestamp between '2016-09-04' and '2018-10-17'
```

## Query results

JOB INFORMATION		RESULTS	CHART
Row	city_count ▼	state_count ▼	
1	4119	27	

**Inference** -The Cities & States of customers who ordered during the given period is 4119 and 27

**2(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,

**COUNT**(**DISTINCT** order\_id) **as** count\_order

**FROM** target\_brazil.orders

**GROUP BY** year

**ORDER BY** year

## Query results

JOB INFORMATION		RESULTS	CHART
Row	year ▼	count_order ▼	
1	2016	329	
2	2017	45101	
3	2018	54011	

**Inference-**By looking at the result of the above query we can easily say that there is a growing trend in an e-commerce business in Brazil. The count of orders has shown an overall upward growing trend with some sort of fluctuations .

**Suggestions-**Although we should consider both count of orders and revenue for the overall business growth there at.

**2(2) 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(DISTINCT order\_id) as no\_of\_orders**

**FROM target\_brazil.orders**

**GROUP BY month**

**ORDER BY month**

## Query results

JOB INFORMATION		RESULTS	CHART
Row	month ▼	no_of_orders ▼	
1	1	8069	
2	2	8508	
3	3	9893	
4	4	9343	
5	5	10573	
6	6	9412	
7	7	10318	
8	8	10843	
9	9	4305	
10	10	4959	

**Inference-**As we can clearly see that yes there is some seasonality in the data of an e-commerce business in Brazil. Talking about the seasonality we can see that there is sudden increase in the no. of orders from Feb to March and also from March to August. Also the count of no. of orders in the month of August is at the peak may be because of some festive season.

**Suggestions-**As we can see seasonal variations through the data mainly during the festive season. So business should plan and strategise their marketing strategies to get maximum sales during this season.

**2(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

SELECT CASE

WHEN EXTRACT(hour FROM o.order\_purchase\_timestamp) BETWEEN 0 and 6 THEN 'Dawn'

WHEN EXTRACT(hour FROM o.order\_purchase\_timestamp) BETWEEN 7 and 12 THEN 'Mornings'

WHEN EXTRACT(hour FROM o.order\_purchase\_timestamp) BETWEEN 13 and 18 THEN  
'Afternoon'

WHEN EXTRACT(hour FROM o.order\_purchase\_timestamp) BETWEEN 19 and 23 THEN 'Night'

END AS hour,

COUNT(o.order\_id) as count\_orders

FROM target\_brazil.orders as o JOIN target\_brazil.customers as c

ON o.customer\_id=c.customer\_id

GROUP BY hour

ORDER BY count\_orders DESC

## Query results

JOB INFORMATION		RESULTS	CHART	JSON
Row	hour ▼	count_orders ▼		
1	Afternoon	38135		
2	Night	28331		
3	Mornings	27733		
4	Dawn	5242		

**Inference-**As we can see from the result of above query that most of the Brazilians place their order in afternoon and night also considering this time to be their leisure time.

**Suggestions-** By understanding the buying patterns of the customers business can allocate proper resources at that peak buying time so that business operations can be optimized.

**3(1) Get the month on month no. of orders placed in each state.**

```
SELECT c.customer_state,  
  
       EXTRACT(month FROM o.order_purchase_timestamp) as month,  
  
       COUNT(o.order_purchase_timestamp) as count_of_orders_  
  
FROM target_brazil.customers as c JOIN target_brazil.orders as o  
  
ON c.customer_id=o.customer_id  
  
GROUP BY c.customer_state,month  
  
ORDER BY c.customer_state,month
```

Query results

JOB INFORMATION		RESULTS	CHART	JSON	EXECUTION D
Row	customer_state	month	count_of_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		

**Inference-**It is clearly showing the customer purchase behaviour monthly on a state by state basis.SP continues to have the largest purchase rate .

**Suggestions-** As we can see that SP has the largest purchase rate as compared to other states so this is the best chance of expanding business in other areas as well so as to get more of the customer's attention .This will increase the orders count and also customers base.

**3(2)How are the customers distributed across all the states?**

```

SELECT customer_state,
COUNT(DISTINCT customer_unique_id) AS total_custmers,
FROM `target_brazil.customers`
GROUP BY customer_state

```



```
ORDER BY total_custmers DESC;
```

## Query results

JOB INFORMATION

RESULTS

CHART

JSON

Row	customer_state	total_custmers
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

Inference-As we can infer from the above query result that state SP has the largest number of customers making it the most populous state. Previously also we have noticed that maximum number of order were from this state only so this establishes a kind of positive correlation between the number of order being placed from this state and the number of people residing at this place.

Suggestions-Getting fair insights of the population of particular place helps in making business strategies to expand business at that particular place. Such insights help in targeting specific regions and allocating resources strategically .

4(1) Get the % increase in the cost of orders from year 2017 to 2018 (include months between Jan to Aug only).

```
SELECT a.months,a.monthly_2017,b. monthly_2018,b.monthly_2018-a.monthly_2017 as
change_in_payments,

ROUND(((b.monthly_2018-a.monthly_2017)/monthly_2017)*100,2) as change_in_percentage

FROM

(SELECT EXTRACT(year FROM order_purchase_timestamp) as years,

EXTRACT(month FROM order_purchase_timestamp) as months,ROUND(SUM(payment_value),0) as
monthly_2017

FROM `target_brazil.orders` o

JOIN `target_brazil.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 years,months

ORDER BY months,years)a

JOIN

(

SELECT EXTRACT(year FROM order_purchase_timestamp) as years,

EXTRACT(month FROM order_purchase_timestamp) as months,ROUND(SUM(payment_value),0) as
monthly_2018
```

```

FROM `target_brazil.orders` o

JOIN `target_brazil.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 years,months

ORDER BY months,years)b

ON a.months = b.months

ORDER BY months

```

Query results							<a href="#">SAVE RESULTS</a>
JOB INFORMATION		RESULTS	CHART	JSON	EXECUTION DETAILS		EXECUTION
Row	months	monthly_2017	monthly_2018	change_in_payments	change_in_percentag		
1	1	138488.0	1115004.0	976516.0	705.13		
2	2	291908.0	992463.0	700555.0	239.99		
3	3	449864.0	1159652.0	709788.0	157.78		
4	4	417788.0	1160785.0	742997.0	177.84		
5	5	592919.0	1153982.0	561063.0	94.63		
6	6	511276.0	1023880.0	512604.0	100.26		
7	7	592383.0	1066541.0	474158.0	80.04		
8	8	674396.0	1022425.0	348029.0	51.61		

**Inference-**As we can see from the result of above query that the percentage increase is more in the month of January followed by February and April.

**4(2) Calculate the Total & Average value of order price for each state.**

```

SELECT

```

```

    c.customer_state,

```

```

ROUND(SUM(o.price)) as total_price,

ROUND(AVG(o.price)) as avg_price

FROM `target_brazil.order_items` as o JOIN `target_brazil.orders` as ord

ON o.order_id=ord.order_id

JOIN `target_brazil.customers` as c ON ord.customer_id=c.customer_id

GROUP BY c.customer_state

ORDER BY total_price DESC, avg_price ASC

```

## Query results

JOB INFORMATION		RESULTS	CHART	JSON	EXECUTION DE
Row	customer_state ▼	total_price ▼	avg_price ▼		
1	SP	5202955.0	110.0		
2	RJ	1824093.0	125.0		
3	MG	1585308.0	121.0		
4	RS	750304.0	120.0		
5	PR	683084.0	119.0		
6	SC	520553.0	125.0		
7	BA	511350.0	135.0		
8	DF	302604.0	126.0		

**Inference-**It is clearly visible that the state SP has the highest total price value and the lowest average price value.

By examining the cost-trends state wise businesses can gain valuable insights in an economic landscapes and can easily able to find opportunities to grow and can make decisions to optimize pricing strategies and enhance logistics.

**4(3)Calculate the Total & Average value of order freight for each state.**

SELECT

```
c.customer_state,  
  
ROUND(SUM(o.freight_value)) as total_freight,  
  
ROUND(AVG(o.freight_value)) as avg_freight  
  
FROM `target_brazil.order_items` as o JOIN `target_brazil.orders` as ord  
  
ON o.order_id=ord.order_id  
  
JOIN `target_brazil.customers` as c ON ord.customer_id=c.customer_id  
  
GROUP BY c.customer_state  
  
ORDER BY total_freight DESC,avg_freight DESC
```

## Query results

JOB INFORMATION		RESULTS	CHART	JSON	EXECUTION
Row	customer_state	total_freight	avg_freight		
1	SP	718723.0	15.0		
2	RJ	305589.0	21.0		
3	MG	270853.0	21.0		
4	RS	135523.0	22.0		
5	PR	117852.0	21.0		
6	BA	100157.0	26.0		
7	SC	89660.0	21.0		
8	PE	59450.0	33.0		

Inference-the data above reveals that the state SP has the highest total freight value but the lowest average freight value, on the other hand state PB has the highest average freight value.

5(1) Find the no. of days taken to deliver each order from the order's purchase date as delivery time.

Also, calculate the difference (in days) between the estimated & actual delivery date of an order.

```
SELECT order_id,

DATE_DIFF(order_delivered_customer_date,order_purchase_timestamp, day)
time_to_deliver,

DATE_DIFF(order_estimated_delivery_date,order_delivered_customer_date, day)
diff_estimated_delivery

FROM `target_brazil.orders`

WHERE order_status='delivered'

ORDER BY order_id
```

Query results

JOB INFORMATION		RESULTS	CHART	JSON	EXECUTIO
Row	order_id	time_to_deliver	diff_estimated_delive		
1	00010242fe8c5a6d1ba2dd792...	7	8		
2	00018f77f2f0320c557190d7a1...	16	2		
3	000229ec398224ef6ca0657da...	7	13		
4	00024acbcd0a6daa1e931b03...	6	5		
5	00042b26cf59d7ce69dfabb4e...	25	15		
6	00048cc3ae777c65dbb7d2a06...	6	14		
7	00054e8431b9d7675808bcb8...	8	16		
8	000576fe39319847cbb9d288c...	5	15		

Inference- Improving delivery time in areas with longer delivery durations helps in gaining a positive impact on customer satisfaction and can increase the purchases. Implementing efficient shipping processes are key to this strategy.

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

```
SELECT c.customer_state,AVG(oi.freight_value) as highest_avg_freight

FROM target_brazil.customers as c JOIN target_brazil.orders as o

ON c.customer_id=o.customer_id JOIN target_brazil.order_items as oi

ON o.order_id=oi.order_id

GROUP BY c.customer_state

ORDER BY highest_avg_freight DESC

LIMIT 5
```

## Query results

JOB INFORMATION		RESULTS	CHART	JSON
Row	customer_state ▼	highest_avg_freight		
1	RR	42.98442307692...		
2	PB	42.72380398671...		
3	RO	41.06971223021...		
4	AC	40.07336956521...		
5	PI	39.14797047970...		

```
SELECT c.customer_state,AVG(oi.freight_value) as lowest_avg_freight

FROM target_brazil.customers as c JOIN target_brazil.orders as o

ON c.customer_id=o.customer_id JOIN target_brazil.order_items as oi
```

```
ON o.order_id=oi.order_id

GROUP BY c.customer_state

ORDER BY lowest_avg_freight ASC
```

LIMIT 5

## Query results

JOB INFORMATION		RESULTS	CHART	JSON
Row	customer_state ▼	lowest_avg_freight		
1	SP	15.14727539041...		
2	PR	20.53165156794...		
3	MG	20.63016680630...		
4	RJ	20.96092393168...		
5	DF	21.04135494596...		

Insights-From the data we can easily conclude that SP has the lowest avg freight value and RR has the highest freight value.

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

```
SELECT DISTINCT c.customer_state,

AVG(DATE_DIFF(o.order_delivered_customer_date ,o.order_purchase_timestamp,DAY))
as highest_avg_delivery_time

FROM target_brazil.customers as c JOIN target_brazil.orders as o

ON c.customer_id=o.customer_id

GROUP BY c.customer_state
```



ORDER BY highest\_avg\_delivery\_time DESC

LIMIT 5

## Query results

JOB INFORMATION		RESULTS	CHART	JSON
Row	customer_state ▼	highest_avg_delivery		
1	RR	28.97560975609...		
2	AP	26.73134328358...		
3	AM	25.98620689655...		
4	AL	24.04030226700...		
5	PA	23.31606765327...		

```
SELECT DISTINCT c.customer_state,
```

```
AVG(DATE_DIFF(o.order_delivered_customer_date ,o.order_purchase_timestamp,DAY)) as  
lowest_avg_delivery_time
```

```
FROM target_brazil.customers as c JOIN target_brazil.orders as o
```

```
ON c.customer_id=o.customer_id
```

```
GROUP BY c.customer_state
```

```
ORDER BY lowest_avg_delivery_time ASC
```

```
LIMIT 5
```

## Query results

JOB INFORMATION		RESULTS	CHART	JSON
Row	customer_state ▼	lowest_avg_delivery		
1	SP	8.298061489072...		
2	PR	11.52671135486...		
3	MG	11.54381329810...		
4	DF	12.50913461538...		
5	SC	14.47956019171...		

Insights-From the results it is clearly visible that state RR has the highest avg delivery time and state SP has the highest delivery time.

5(4)Find out the top 5 states where the order delivery is really fast as compared to the estimated date of delivery.

```
SELECT DISTINCT c.customer_state,

               AVG(DATE_DIFF(o.order_estimated_delivery_date,
o.order_delivered_customer_date,DAY))

               OVER (PARTITION BY c.customer_state) as top_5_fastest_delivery

FROM `target_brazil.orders` AS o

JOIN `target_brazil.customers` AS c

ON o.customer_id = c.customer_id

ORDER BY top_5_fastest_delivery

LIMIT 5;
```

## Query results

JOB INFORMATION

RESULTS

CHART

JS

Row	customer_state ▼	top_5_fastest_delivery
1	AL	7.947103274559...
2	MA	8.768479776847...
3	SE	9.173134328358...
4	ES	9.618546365914...
5	BA	9.934889434889...

Inference-Above results shows the fastest delivery time in 5 states.

6(1) Find the month on month no. of orders placed using different payment types.

```
SELECT p.payment_type,  
       EXTRACT(month FROM o.order_purchase_timestamp) as mnth,  
       COUNT(DISTINCT o.order_id) as no_of_orders FROM `target_brazil.orders` as o  
JOIN `target_brazil.payments` as p ON o.order_id=p.order_id  
GROUP BY p.payment_type,mnth  
ORDER BY p.payment_type,mnth
```

## Query results

JOB INFORMATION		RESULTS	CHART	JSON	EXECUTION
Row	payment_type ▼	mnth ▼	no_of_orders ▼		
1	UPI	1	1715		
2	UPI	2	1723		
3	UPI	3	1942		
4	UPI	4	1783		
5	UPI	5	2035		
6	UPI	6	1807		
7	UPI	7	2074		
8	UPI	8	2077		

## Query results

JOB INFORMATION		RESULTS	CHART	JSON	EXECUTION DE
Row	payment_type ▼	mnth ▼	no_of_orders ▼		
9	UPI	9	903		
10	UPI	10	1056		
11	UPI	11	1509		
12	UPI	12	1160		
13	credit_card	1	6093		
14	credit_card	2	6582		
15	credit_card	3	7682		
16	credit_card	4	7276		

**Insights-**The analysis shows and upward trend from Jan to Aug and another upward trend from Sep to Nov.Credit card users are the most followed by UPI transactions.One possible reason for increase in credit card users may be because of the feature of buy now and pay later feature.

6(2) Find the no. of orders placed on the basis of the payment installments that have been paid.

```
SELECT DISTINCT payment_installments,  
COUNT(order_id) OVER(PARTITION BY payment_installments) AS count_of_orders_  
FROM target_brazil.payments  
ORDER BY 1,2
```

## Query results

JOB INFORMATION		RESULTS	CHART
Row	payment_installment	count_of_orders_	
1	0	2	
2	1	52546	
3	2	12413	
4	3	10461	
5	4	7098	
6	5	5239	
7	6	3920	
8	7	1626	

Insights- The majority of orders has the payment installment of 1 .

### Actionable Insights-

1.As we have seen that state of SP has more number of orders than the other states which means there is a good opportunity to expand businesses in all the remaining area so that there will be more customers and business will also expand.

2.Improving delivery time in areas with longer delivery times can have positive impact on customers which in turn will help in customer retention.

3.States like SP and RJ already have high order counts.So there is need to boost sales in other area through loyalty programmes,personalized marketing campaigns.

**4.Analyzing impact of economic conditions on sales can help us identifying areas of improvement .**

**5.As it is inferenced from the above results that in some of the months sales have got reduced.So for compensating with such situations it is good to provide discounts on the products,so that customers purchases increases during off season too.**

**6.AS there are seasonal variations so business should plan their marketing and sales strategy so as to have more profits during the peak season.**

#### **Recommendations-**

**1.Enhance the customer experience by providing some chat support services so that their queries and problems can get resolved at the earliest.**

**2.Always have an eye over the competitor's activity and business so that we can strategise in a best way possible with respect to pricing .**

**3.Collaboration with the sellers is a way to have more variety in food and other things and also improve product quality.**

**4.Encourage customer retention staregies so that we can have repeat purchases.This can be done through loyalty programs and offering some referral awards.**

**5.Ensure price and freight value to have competitiveness in the market while maximizing profit and revenues.**