

Business Case Study: Target SQL



Details:

Name : Tanmoy Basuli

Email-Id : basuli575tanmoy@gmail.com

Context:

Target is a globally renowned brand and a prominent retailer in the United States. Target makes itself a preferred shopping destination by offering outstanding value, inspiration, innovation, and an exceptional guest experience that no other retailer can deliver.

This business case focuses on the operations of Target in Brazil and provides insightful information about 100,000 orders placed between 2016 and 2018. The dataset offers a comprehensive view of various dimensions including the order status, price, payment and freight performance, customer location, product attributes, and customer reviews.

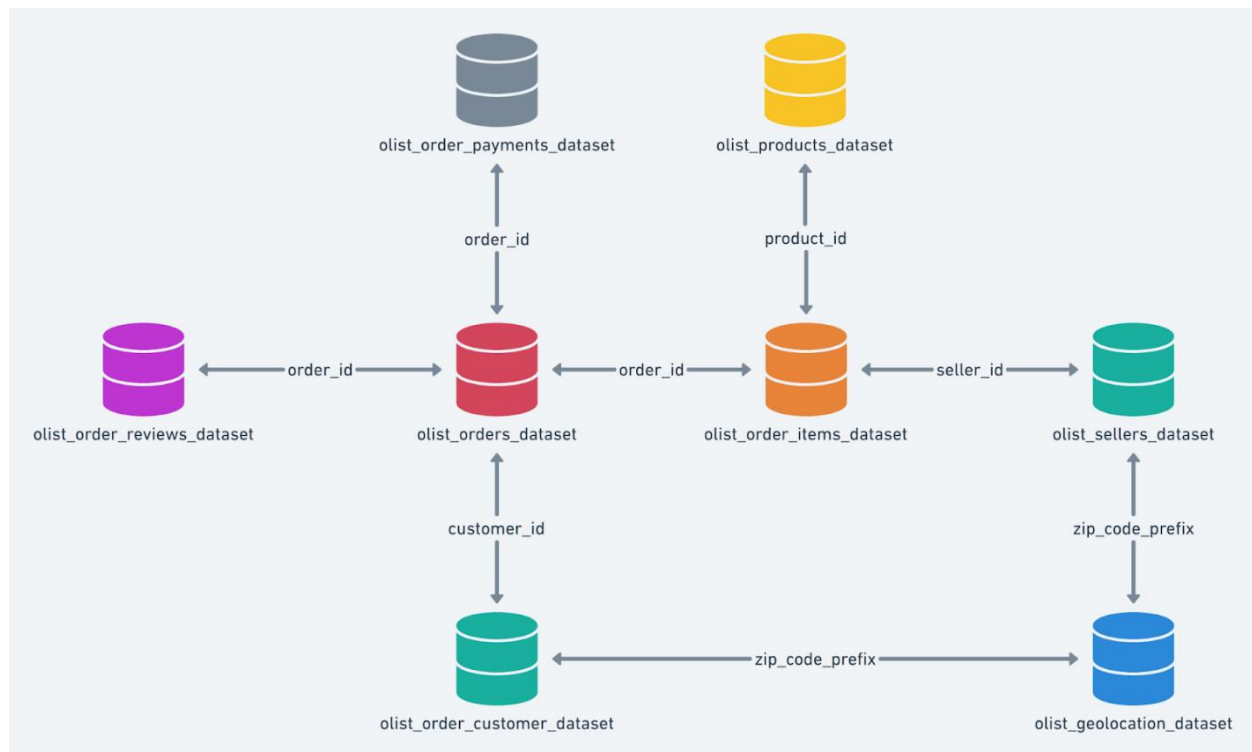
By analyzing this extensive dataset, it becomes possible to gain valuable insights into Target's operations in Brazil. The information can shed light on various aspects of the business, such as order processing, pricing strategies, payment and shipping efficiency, customer demographics, product characteristics, and customer satisfaction levels.

Dataset: <https://drive.google.com/drive/folders/1TGEc66YKbD443nslRi1bWgVd238gJCnb>

The data is available in 8 csv files:

1. customers.csv
2. sellers.csv
3. order_items.csv
4. geolocation.csv
5. payments.csv
6. reviews.csv
7. orders.csv
8. products.csv

Dataset schema:



The column description for these csv files is given below.

The **customers.csv** contain following features:

Features	Description
customer_id	ID of the consumer who made the purchase
customer_unique_id	Unique ID of the consumer
customer_zip_code_prefix	Zip Code of consumer's location
customer_city	Name of the City from where order is made
customer_state	State Code from where order is made (Eg. são paulo - SP)

The **sellers.csv** contains following features:

Features	Description
seller_id	Unique ID of the seller registered
seller_zip_code_prefix	Zip Code of the seller's location
seller_city	Name of the City of the seller

seller_state State Code (Eg. são paulo - SP)

The **order_items.csv** contain following features:

Features	Description
order_id	A Unique ID of order made by the consumers
order_item_id	A Unique ID given to each item ordered in the order
product_id	A Unique ID given to each product available on the site
seller_id	Unique ID of the seller registered in Target
shipping_limit_date	The date before which the ordered product must be shipped
price	Actual price of the products ordered
freight_value	Price rate at which a product is delivered from one point to another

The **geolocations.csv** contain following features:

Features	Description
geolocation_zip_code_prefix	First 5 digits of Zip Code
geolocation_lat	Latitude
geolocation_lng	Longitude
geolocation_city	City
geolocation_state	State

The **payments.csv** contain following features:

Features	Description
order_id	A Unique ID of order made by the consumers
payment_sequential	Sequences of the payments made in case of EMI
payment_type	Mode of payment used (Eg. Credit Card)
payment_installments	Number of installments in case of EMI purchase
payment_value	Total amount paid for the purchase order

The **orders.csv** contain following features:

Features	Description
order_id	A Unique ID of order made by the consumers
customer_id	ID of the consumer who made the purchase
order_status	Status of the order made i.e. delivered, shipped, etc.

order_purchase_timestamp	Timestamp of the purchase
order_delivered_carrier_date	Delivery date at which carrier made the delivery
order_delivered_customer_date	Date at which customer got the product
order_estimated_delivery_date	Estimated delivery date of the products

The **reviews.csv** contain following features:

Features	Description
review_id	ID of the review given on the product ordered by the order id
order_id	A Unique ID of order made by the consumers
review_score	Review score given by the customer for each order on a scale of 1-5
review_comment_title	Title of the review
review_comment_message	Review comments posted by the consumer for each order
review_creation_date	Timestamp of the review when it is created
review_answer_timestamp	Timestamp of the review answered

The **products.csv** contain following features:

Features	Description
product_id	A Unique identifier for the proposed project.
product_category_name	Name of the product category
product_name_lenght	Length of the string which specifies the name given to the products on the site
product_description_lenght	Length of the description written for each product ordered on the site
product_photos_qty	Number of photos of each product ordered available on the shopping site
product_weight_g	Weight of the products ordered in grams
product_length_cm	Length of the products ordered in centimeters
product_height_cm	Height of the products ordered in centimeters
product_width_cm	Width of the product ordered in centimeters

Assume you are a data scientist at Target and are given this data to analyze and provide some insights and recommendations from it.

What does 'good' look like?

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.
2. Get the time range between which the orders were placed.
3. Count the Cities & States of customers who ordered during the given period.
2. **In-depth Exploration:**
 1. Is there a growing trend in the no. of orders placed over the past years?
 2. Can we see some kind of monthly seasonality in terms of the no. of orders being placed?
 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
3. **Evolution of E-commerce orders in the Brazil region:**
 1. Get the month on month no. of orders placed in each state.
 2. How are the customers distributed across all the states?
4. **Impact on Economy: Analyze the money movement by e-commerce by looking at order prices, freight and others.**
 1. Get the % increase in the cost of orders from year 2017 to 2018 (include months between Jan to Aug only).
You can use the "payment_value" column in the payments table to get the cost of orders.
 2. Calculate the Total & Average value of order price for each state.
 3. Calculate the Total & Average value of order freight for each state.
5. **Analysis based on sales, freight and delivery time.**
 1. Find the no. of days taken to deliver each order from the order's purchase date as delivery time.
Also, calculate the difference (in days) between the estimated & actual delivery date of an order.
Do this in a single query.

You can calculate the delivery time and the difference between the estimated & actual delivery date using the given formula:
 - **time_to_deliver** = order_delivered_customer_date - order_purchase_timestamp
 - **diff_estimated_delivery** = order_delivered_customer_date - order_estimated_delivery_date
 2. Find out the top 5 states with the highest & lowest average freight value.
 3. Find out the top 5 states with the highest & lowest average delivery time.
 4. Find out the top 5 states where the order delivery is really fast as compared to the estimated date of delivery.

You can use the difference between the averages of actual & estimated delivery date to figure out how fast the delivery was for each state.

6. Analysis based on the payments:

1. Find the month on month no. of orders placed using different payment types.
2. Find the no. of orders placed on the basis of the payment installments that have been paid.

Now try to analyze the data :

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.

Filter Enter property name or value ?								
<input type="checkbox"/>	Field name	Type	Mode	Key	Collation	Default Value	Policy Tags ?	Description
<input type="checkbox"/>	customer_id	STRING	NULLABLE					
<input type="checkbox"/>	customer_unique_id	STRING	NULLABLE					
<input type="checkbox"/>	customer_zip_code_prefix	INTEGER	NULLABLE					
<input type="checkbox"/>	customer_city	STRING	NULLABLE					
<input type="checkbox"/>	customer_state	STRING	NULLABLE					


INSIGHTS:

In Big Query section clicked on customers table and all the details came up with above result.

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

```
select
min(order_purchase_timestamp) as Started_Order_Time,
max(order_purchase_timestamp) as Last_Ordered_Time
from Target_Info.orders;
```

Query results

 SAVE RE

JOB INFORMATIONRESULTSJSONEXECUTION DETAILSCHARTPREVIEW

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

INSIGHTS:

This result allowed us to find the start and end dates of the orders range where customers made order i.e. from 4th September 2016 to 17th October 2018.

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

QUERY:

```
select
distinct c.customer_city,
c.customer_state,
count(o.order_id) as customer_count
from Target_Info.customers c join Target_Info.orders o using(customer_id)
group by 1,2
order by customer_count desc;
```

Query results

[SAVE RESULT](#)

JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS	CHART	PREVIEW	EXEC
Row	customer_city ▼	customer_state ▼	customer_count ▼			
1	sao paulo	SP	15540			
2	rio de janeiro	RJ	6882			
3	belo horizonte	MG	2773			
4	brasilgia	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			

INSIGHTS:

Here, we can see that sao paulo city from SP state alone has more orders than the following cities. Customers from all above listed output cities and states Ordered in the given time.

2. In-depth Exploration:

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

QUERY:

```
select
Extract(Year from order_purchase_timestamp) as Year,
Extract(Month from order_purchase_timestamp) as Month,
count( distinct o.order_id) as count_order
from Target_Info.orders o join Target_Info.customers c using(customer_id)
group by 1,2
order by 1,2;
```

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	CHART	PREVIEW
Row	Year	Month	count_order			
1	2016	9	4			
2	2016	10	324			
3	2016	12	1			
4	2017	1	800			
5	2017	2	1780			
6	2017	3	2682			
7	2017	4	2404			
8	2017	5	3700			
9	2017	6	3245			
10	2017	7	4026			

INSIGHTS:

Result came, based on each month of each year. Based on the analysis of order count, it can be observed that there is some up and down in the year of 2016 but can see there is a growing trend from the year 2017 in e-commerce in Brazil. The count of purchases has shown an overall upward trend, with some fluctuations.

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

QUERY:

```
select
extract(month from order_purchase_timestamp) as Month,
count( distinct o.order_id) as count_order
from Target_Info.orders o join Target_Info.customers c using(customer_id)
group by 1
order by 1;
```

Query results

[SAVE RI](#)

JOB INFORMATION		RESULTS		JSON	EXECUTION DETAILS	CHART	PREVIEW
Row	Month		count_order				
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				

INSIGHTS:

Considering this dataset and this result above, it's quite look like trend somewhere can be draw level and somewhere up and down regarding seasonality trends. However, from the analysis we can observe some seasonality in the e-commerce orders. The count of orders generally increases from April to August with fluctuations in between. Notably, there is an increase in orders during February and March. Additionally, the month of August shows a peak in order count and in the month of September there is a less count.

3. During what time of the day, do the Brazilian customers mostly place their orders?
(Dawn, Morning, Afternoon or Night)
1. 0-6 hrs : Dawn
 2. 7-12 hrs : Mornings
 3. 13-18 hrs : Afternoon
 4. 19-23 hrs : Night

QUERY:

```
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 Session,
count( distinct o.order_id) as count_order
from Target_Info.orders o join Target_Info.customers c using(customer_id)
group by 1
order by 2 desc;
```

Query results

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	CHART	PREVIEW
Row	Session	count_order				
1	Afternoon	38135				
2	Night	28331				
3	Mornings	27733				
4	Dawn	5242				

INSIGHTS:

Based on the analysis after output result, we found that Brazilian customers tend to place most orders during the daytime, specifically in the afternoon. Night and Morning following by the Afternoon. We can say that those customers prefer to order once they have free and relax time or we can say after returning from the office they enjoy the order. In early morning they don't show much interest to order more.

3. Evolution of E-commerce orders in the Brazil region:

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

QUERY:

```
select
extract(month from order_purchase_timestamp) as Month,
c.customer_state,
count( distinct o.order_id) as count_order
from Target_Info.orders o join Target_Info.customers c using(customer_id)
group by 2,1
order by 2,1;
```

Query results

[SAVE RESULTS](#)

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	CHART	PREVIEW	EXEC
Row	Month	customer_state	count_order				
1	1	AC	8				
2	2	AC	6				
3	3	AC	4				
4	4	AC	9				
5	5	AC	10				
6	6	AC	7				
7	7	AC	9				
8	8	AC	7				
9	9	AC	5				
10	10	AC	6				

INSIGHTS:

The above result described the month-on-month order counts in each state of Brazil, providing valuable insights into the customer purchase trends on a state-by-state basis. There are some kind of similar count of ordered in each month but if we are going below output result (SP) consistently has the highest number of orders in any given month, followed by (RJ) and (MG).

- How are the customers distributed across all the states?

QUERY:

```
select
customer_state,
count(customer_id) as Number_Of_Customers
from Target_Info.customers
group by 1
order by 2 desc;
```

Query results



JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	CHART	PREVIEW
Row	customer_state ▼	Number_Of_Customers				
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				

INSIGHTS:

The data reveals that the state of (SP) has the highest number of customers, which can be attributed to its status as the most populous state in Brazil.

4. Impact on Economy: Analyze the money movement by e-commerce by looking at order prices, freight and others.

1. Get the % increase in the cost of orders from year 2017 to 2018 (include months between Jan to Aug only).
You can use the "payment_value" column in the payments table to get the cost of orders.

QUERY:

```
select
extract(month from order_purchase_timestamp) as Month_Number_Wisee,
FORMAT_DATE('%B', order_purchase_timestamp) as Month_Name_Wise,
Round
(
  (
    (
      sum(case when extract(year from order_purchase_timestamp) = 2018 and extract(month
from order_purchase_timestamp) between 1
and 8 then p.payment_value end)
    -
      sum(case when extract(year from order_purchase_timestamp) = 2017 and extract(month
from order_purchase_timestamp) between 1
and 8 then p.payment_value end)
    )
  /
    sum(case when extract(year from order_purchase_timestamp) = 2017 and extract(month
from order_purchase_timestamp) between 1
and 8 then p.payment_value end)
  ) * 100 ,2
) as percentage_increase
from Target_Info.orders o join Target_Info.payments p using(order_id)
where extract(year from order_purchase_timestamp) IN(2017,2018)
AND extract(month from order_purchase_timestamp) between 1 and 8
group by 1,2
order by 1;
```

Query results

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	CHART	PREV
Row	Month_Number_Wise	Month_Name_Wise	percentage_increase			
1	1	January	705.13			
2	2	February	239.99			
3	3	March	157.78			
4	4	April	177.84			
5	5	May	94.63			
6	6	June	100.26			
7	7	July	80.04			
8	8	August	51.61			

INSIGHTS:

To better view, showed the month of number along with the appropriate month name with some percentage increased in month wise. Upon on the above analysis the overall percentage increase in the cost of orders from 2017 to 2018, including only the months from January to August, is 138.53%. Upon examining the month-wise increase, January shows the highest percentage increase, is 705.13 followed by February and April. August is showing the minimum number of increase as compared to 2016 and 2017.

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

QUERY:

```
select
distinct c.customer_state,
round(sum(p.price),2) as Sum_Of_Orders_Price,
```



```

round(avg(p.price),2) as Avg_Of_Orders_Price
from Target_Info.customers c join Target_Info.orders o using(customer_id)
join Target_Info.order_items p using(order_id)
group by 1;

```

Query results



JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	CHART	PREVIEW
Row	customer_state	Sum_Of_Orders_Price	Avg_Of_Orders_Price			
1	MT	156453.53	148.3			
2	MA	119648.22	145.2			
3	AL	80314.81	180.89			
4	SP	5202955.05	109.65			
5	MG	1585308.03	120.75			
6	PE	262788.03	145.51			
7	RJ	1824092.67	125.12			
8	DF	302603.94	125.77			
9	RS	750304.02	120.34			
10	SE	58920.85	153.04			

INSIGHTS:

The analysis reveals interesting findings. While (SP) has the highest total price value, it surprisingly has the lowest average price value among all states. On the other hand, the state of (PB) has the highest average price value.

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

QUERY:

```

select
distinct c.customer_state,
round(sum(p.freight_value),2) as Sum_Of_Orders_Freight,

```

```

round(avg(p.freight_value),2) as Avg_Of_Orders_Freight
from Target_Info.customers c join Target_Info.orders o using(customer_id)
join Target_Info.order_items p using(order_id)
group by 1;

```

Query results

[SAVE RESULTS](#)

JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS	CHART	PREVIEW	EXECU
row	customer_state ▼	Sum_Of_Orders_Freight ▼	Avg_Of_Orders_Freight ▼			
1	MT	29715.43	28.17			
2	MA	31523.77	38.26			
3	AL	15914.59	35.84			
4	SP	718723.07	15.15			
5	MG	270853.46	20.63			
6	PE	59449.66	32.92			
7	RJ	305589.31	20.96			
8	DF	50625.5	21.04			
9	RS	135522.74	21.74			
10	SE	14111.47	36.65			

INSIGHTS:

The analysis reveals interesting findings. While (SP) has the highest Orders freight value and, it surprisingly has the lowest average freight value among all states. On the other hand, the state of (PB) has the highest average freight value.

5. Analysis based on sales, freight and delivery time.

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

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

Do this in a single query.

You can calculate the delivery time and the difference between the estimated & actual delivery date using the given formula:

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

QUERY:

SELECT

```
order_id,
DATE_DIFF(order_delivered_customer_date, order_purchase_timestamp, DAY)
AS Delivered_in_days,
DATE_DIFF(order_estimated_delivery_date, order_delivered_customer_date, DAY)
AS Estimated_minus_actual_delivery_days
from Target_Info.orders
WHERE
DATE_DIFF(order_delivered_customer_date, order_purchase_timestamp, DAY) IS NOT NULL
ORDER BY 2;
```

Query results

[SAVE RE](#)

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	CHART	PREVIEW
Row	order_id	Delivered_in_days	Estimated_minus_ac			
8	38c1e3d4ed6a13cd0cf612d4c09766e9	0	16			
9	8339b608be0d84fca9d8da68b58332c3	0	27			
10	f349cdb62f69c3fae5c4d7d3f3a4a185	0	12			
11	f3c6775ba3d2d9fe2826f93b71f12008	0	11			
12	b70a8d75313560b4acf607739560a0e5	0	9			
13	21a8ffca665bc7a1087d31751a7b7cbc	0	11			
14	44558a1547e448b41c48c4087fe32ddd	1	5			
15	3bfd703ce884b8a0a65e63f2d4db6393	1	5			
16	68fa625f02107978e969340da28d0c89	1	5			
17	0d8f485ffe96c81fe3e282095e942c2e	1	11			

INSIGHTS:

Based on the requirement showed time to deliver and difference estimated delivery on each different order.

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

QUERY:

```
select customer_state,Avg_Freight_Value,heightst_rnk as Rank
from
(
select
c.customer_state,round(avg(p.freight_value),2) as Avg_Freight_Value,
dense_rank() over(order by avg(p.freight_value) desc) as heightst_rnk,
dense_rank() over(order by avg(p.freight_value) asc) as lowest_rnk
from Target_Info.customers c join Target_Info.orders o using(customer_id)
join Target_Info.order_items p using(order_id)
group by 1
) t
where t.heightst_rnk <=5 OR lowest_rnk <=5
order by 2 desc
```

Query results

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	CHART	PRE
Row	customer_state	Avg_Freight_Value	Rank			
1	RR	42.98	1			
2	PB	42.72	2			
3	RO	41.07	3			
4	AC	40.07	4			
5	PI	39.15	5			
6	DF	21.04	23			
7	RJ	20.96	24			
8	MG	20.63	25			
9	PR	20.53	26			
10	SP	15.15	27			

INSIGHTS:

Based on the requirement displayed top 5 highest & lowest average freight value. In the output table we can easily notice the state (RR) has the highest avg freight value followed by state PB, RO, AC and PI.

Similarly showed states last 5 lowest avg freight value. States are DF, RJ, MG, PR, SP. Showed rank number in the last. So, we can easily find out the requirement.

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

QUERY:

```
select customer_state, avg_time_to_delivery, heightst_rnk as Rank
from
(
select
c.customer_state,
ROUND(AVG(DATE_DIFF(order_delivered_customer_date, order_purchase_timestamp, DAY)), 2)
  AS avg_time_to_delivery,
dense_rank() over(order by AVG(DATE_DIFF(order_delivered_customer_date,
order_purchase_timestamp, DAY)) desc) as heightst_rnk,
dense_rank() over(order by AVG(DATE_DIFF(order_delivered_customer_date,
order_purchase_timestamp, DAY)) asc) as lowest_rnk
from Target_Info.customers c join Target_Info.orders o using(customer_id)
where DATE_DIFF(order_delivered_customer_date, order_purchase_timestamp, DAY) is not
null
group by 1
) t
where t.heightst_rnk <=5 OR lowest_rnk <=5
order by 2 desc;
```

Query results

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	CHART	PREVIEW
Row	customer_state ▼	avg_time_to_delivery	Rank ▼			
1	RR	28.98	1			
2	AP	26.73	2			
3	AM	25.99	3			
4	AL	24.04	4			
5	PA	23.32	5			
6	SC	14.48	23			
7	DF	12.51	24			
8	MG	11.54	25			
9	PR	11.53	26			
10	SP	8.3	27			

INSIGHTS:

Based on the requirement displayed top 5 highest & lowest average delivery time. In the output table we can easily notice the state (RR) has the highest average delivery time followed by state AP, AM, AL and PA.

Similarly showed states last 5 lowest average delivery time. States are SC, DF, MG, PR and SP. Showed rank number in the last. So, we can easily find out the requirement.

4. Find out the top 5 states where the order delivery is really fast as compared to the estimated date of delivery.
You can use the difference between the averages of actual & estimated delivery date to figure out how fast the delivery was for each state.

QUERY:

```
select customer_state, Fast_delivery
from
(
select
c.customer_state,
ROUND(AVG(DATE_DIFF(o.order_estimated_delivery_date, o.order_delivered_customer_date,
DAY))), 2)
AS Fast_delivery,
dense_rank() over(order by AVG(DATE_DIFF(o.order_estimated_delivery_date,
o.order_delivered_customer_date, DAY)) desc) as heightst_rnk
from Target_Info.customers c join Target_Info.orders o using(customer_id)
group by 1
) t
where t.heightst_rnk <=5
order by 2 desc
```

Query results

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS
Row	customer_state	Fast_delivery		
1	AC	19.76		
2	RO	19.13		
3	AP	18.73		
4	AM	18.61		
5	RR	16.41		

INSIGHTS:

Based on the requirement displayed top 5 states where the order delivery is really fast as compared to the estimated date of delivery. In the output table we can easily notice the state (AC) has quick delivery time followed by state RO, AP, AM and RR.

6. Analysis based on the payments:

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

QUERY:

```
select
extract(month from order_purchase_timestamp) as Month,
p.payment_type,
count(distinct o.order_id) as No_Of_Order_Placed
from Target_Info.orders o join Target_Info.payments p using(order_id)
group by 2,1
order by 2,1;
```

Query results



JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	CHART	PREVIEW
Row	Month	payment_type	No_Of_Order_Placed			
1	1	UPI	1715			
2	2	UPI	1723			
3	3	UPI	1942			
4	4	UPI	1783			
5	5	UPI	2035			
6	6	UPI	1807			
7	7	UPI	2074			
8	8	UPI	2077			
9	9	UPI	903			
10	10	UPI	1056			

Query results

[SA](#)

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	CHART	PREVIEW
Row	Month	payment_type	No_Of_Order_Placed			
11	11	UPI	1509			
12	12	UPI	1160			
13	1	credit_card	6093			
14	2	credit_card	6582			
15	3	credit_card	7682			
16	4	credit_card	7276			
17	5	credit_card	8308			
18	6	credit_card	7248			
19	7	credit_card	7810			
20	8	credit_card	8235			

INSIGHTS:

The analysis shows an overall uptrend from January to August and another uptrend from September to November. Credit card transactions are the most popular payment method, followed by UPI. Debit card transactions are the least preferred option. Notably, credit card transactions are rapidly increasing compared to other payment methods, possibly due to benefits like “buy now, pay later” options or cashback received using credit cards.

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

QUERY:

```
select
p.payment_installments,
count(distinct o.order_id) as No_Of_Order_Placed
from Target_Info.orders o join Target_Info.payments p using(order_id)
where o.order_status != 'canceled'
group by 1
order by 2 desc;
```

Query results

JOB INFORMATION		RESULTS	JSON	EXECUTION DETA
Row	payment_installment	No_Of_Order_Placed		
1	1	48732		
2	2	12329		
3	3	10374		
4	4	7046		
5	10	5279		
6	5	5204		
7	8	4224		
8	6	3894		
9	7	1617		
10	9	638		

INSIGHTS:

The analysis reveals that most of the orders (maximum count) have only one payment installment. The highest number of installments is 24, which is associated with 18 orders, and you can see the output above of the result set.

Actionable Insights

1. The data reveals that the state of SP has significantly more orders than the next five states combined. This indicates an opportunity for improvement in the other states. Focusing on these states can help increase the number of orders and expand the customer base.
2. Seasonal variations in sales are observed, with increased sales during festive periods. Businesses should plan their marketing and sales strategies accordingly to capitalize on these peak periods and enhance customer satisfaction, resulting in overall sales growth.
3. Improving delivery times in areas with longer delivery durations can have a positive impact on customer satisfaction and encourage repeat purchases. Streamlining logistics and implementing efficient shipping processes are key to achieving this.
4. States like SP and RJ already have high order counts. To further boost sales and foster brand loyalty, it is recommended to focus on customer retention strategies, such as personalized marketing campaigns, loyalty programs, and exceptional customer service experiences.
5. Analyzing customer demographics can provide valuable insights for tailoring products and marketing strategies to specific target audiences. This customization can lead to increased sales and customer satisfaction.
6. The data indicates a decline in orders during September and October. Offering discounts or promotions during off-peak seasons can incentivize customers to make purchases during these periods, thus boosting sales.

7. While the data does not include information on economic conditions, analyzing their impact on sales can help identify areas for improvement and investment, ensuring resilience during economic fluctuations.

Recommendations

1. Improve logistics and shipping processes to reduce delivery times and enhance customer satisfaction. This includes optimizing warehouse operations, refining shipping routes, and partnering with reliable courier services.
2. Implement customer retention strategies to encourage repeat purchases and foster loyalty. This can be achieved through loyalty programs, referral rewards, and personalized offers.
3. Evaluate pricing and freight fees to ensure competitiveness in the market while maximizing revenue and profitability. Consider increasing prices or adjusting freight fees as appropriate.
4. Invest in technology and infrastructure to enhance the e-commerce experience. This includes implementing chatbots for customer support, improving website performance, and offering personalized product recommendations based on customer behavior.
5. Collaborate with sellers to expand product offerings and improve product quality, catering to diverse customer needs and preferences.
6. Leverage social media platforms and influencers to promote products and increase brand awareness, as they have a strong influence on purchasing decisions in Brazil.

7. Enhance the customer service experience by offering chat support services and ensuring prompt and effective responses to customer inquiries.
8. Monitor competitor activity and adjust the business strategy accordingly, such as matching or offering better pricing, expanding product offerings, or improving customer service to stay competitive in the market.

By implementing these actionable insights and recommendations, businesses can optimize their operations, enhance customer satisfaction, and drive overall sales growth in the Brazilian e-commerce market.

Conclusion

In conclusion, the analysis of e-commerce data in the Brazilian market provides valuable insights into customer buying patterns, sales trends, payment preferences, and delivery experiences. By understanding these patterns and trends, businesses can make informed decisions and implement strategies to optimize their operations and drive growth. Here are the key takeaways from the analysis:

Key Takeaways

- The state of SP dominates the e-commerce market in Brazil, indicating the need to focus on other states for potential growth opportunities.
- Analyzing customer demographics can help tailor products and marketing strategies to specific target audiences, leading to increased sales.

- Offering discounts during off-peak seasons can incentivize customers and boost sales during slower periods.