TARGET CORPORATION

Business Case Study

I. Import the dataset and do usual exploratory analysis steps like checking the structure & characteristics of the dataset:

A. Data type of all columns in the "customers" table.

QUERY:

```
select column_name, data_type
from
`scaler-dsml sequal.E_COMMERCE_TARGET.INFORMATION_SCHEMA.COLUMNS`
where table_name = 'Customers';
```

OUTPUT:

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

INSIGHTS:

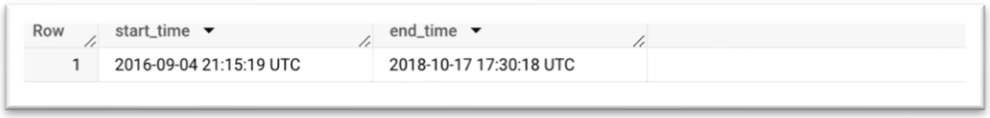
The above table provides an overview of the data types for each column in the “Customers” table of the Scaler-dsml_Sequal E_commerce_target information, which can be useful for data analysis and further processing.

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

QUERY:

```
select
min(order_purchase_timestamp) as start_time,
max(order_purchase_timestamp) as end_time
from `E_COMMERCE_TARGET.Orders`;
```

OUTPUT:



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

ASSUMPTIONS:

- **Start Time:** The minimum order_purchase_timestamp indicates the earliest/start time when an order was placed in the E-commerce system. This timestamp is essential to identify the beginning of the analyzed time range.
- **End Time:** The maximum order_purchase_timestamp represents the latest time when an order was placed in the E-commerce system. This timestamp is crucial for understanding the end of the analyzed time range.

INSIGHTS:

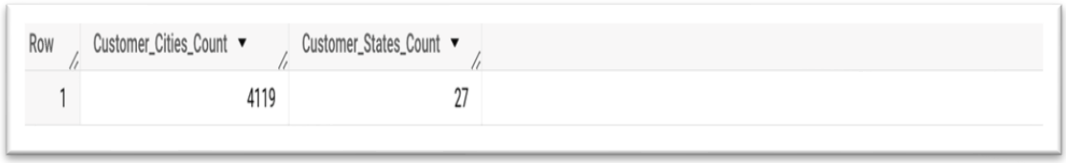
- **Time Range:** The time range between the earliest and latest times provides a comprehensive understanding of the duration during which orders were placed on the E-commerce platform.

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

QUERY:

```
SELECT
Count (distinct customer_city) as Customer_Cities_Count,
Count (distinct customer_state) as Customer_States_Count
from `E_COMMERCE_TARGET.Customers` A
join `E_COMMERCE_TARGET.Orders` B
on A.customer_id = B.customer_id;
```

OUTPUT:



Row	Customer_Cities_Count	Customer_States_Count
1	4119	27

INSIGHTS:

The above table provides the insights on the number of unique customer cities and states who ordered in the given period i.e., there are in total 4119 unique customer cities and 27 unique states in the "Customers" table. This insight helps in understanding the overall market penetration and potential for growth in various regions.

II. In-depth Exploration:

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

QUERY:

```
SELECT
Count (order_id) no_of_orders,
format_date ('%Y-%m',order_purchase_timestamp) as order_month_wise
from `E_COMMERCE_TARGET.Orders`
group by order_month_wise
order by order_month_wise;
```

OUTPUT:

Row	no_of_orders	order_month_wise
1	4	2016-09
2	324	2016-10
3	1	2016-12
4	800	2017-01
5	1780	2017-02
6	2682	2017-03
7	2404	2017-04
8	3700	2017-05
9	3245	2017-06
10	4026	2017-07
11	4331	2017-08
12	4285	2017-09
13	4631	2017-10
14	7544	2017-11
15	5673	2017-12

INSIGHTS:

The above table helps us to analyze the sales trends over the past years monthly. By examining the order count for each month, we can identify peak seasons are during 11th month of 2017 and 3rd month of 2018, slow periods are during 12th month of 2016, 1st month of 2017 and 10th month of 2018.

RECOMMENDATIONS:

To increase the sales, establish real-time monitoring of order counts to identify sudden shifts in demand or potential anomalies that may require immediate action.

B. 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 order_month,
Count (*) as Monthly_order_count
FROM `E_COMMERCE_TARGET.Orders`
GROUP BY order_month
ORDER BY order_month,Monthly_order_count desc;
```

OUTPUT:

Row	order_month	Monthly_order_count
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
11	11	7544
12	12	5674

INSIGHTS:

The above table provides insights into which months have the highest monthly order count. From the above table we can predict the monthly order count is higher in the summer and monsoon seasons I.e., higher during 8^{th}, 7^{th}, 5^{th}, 6^{th} months.

RECOMMENDATIONS:

Based on the insights gained from popular and low-performing months, devise seasonal promotions and marketing strategies to capitalize on peak periods and stimulate demand during slow months. Stock up on popular products before peak months to meet customer demand effectively and avoid overstocking during slower months.

C. 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:

```
SELECT
case
when Extract (Hour from order_purchase_timestamp) between 0 and 6
then 'Dawn'
when Extract (Hour from order_purchase_timestamp) between 7 and 12
then 'Morning'
when Extract (Hour from order_purchase_timestamp) between 13 and 18
then 'Afternoon'
when Extract (Hour from order_purchase_timestamp) between 19 and 23
then 'Night'
End as Time_of_day,
Count (*) as order_count
from `E_COMMERCE_TARGET.Orders`
group by Time_of_day
order by order_count;
```

OUTPUT:

Row	Time_of_day	order_count
1	Dawn	5242
2	Morning	27733
3	Night	28331
4	Afternoon	38135

INSIGHTS:

- By observing the above table, we can see that the peak order activity is during afternoons. This period might correspond to the time when customers are most active and engaged in online shopping activities.
- Night and Morning recorded similar order counts, with 28,331 and 27,733 orders, respectively. This indicates a consistent level of customer activity during these periods.
- The Dawn period saw 5,242 orders, which is notably lower compared to other time categories.

RECOMMENDATIONS:

'TARGET' can consider implementing targeted marketing strategies or exclusive early morning offers to attract more customers during this time and boost sales.

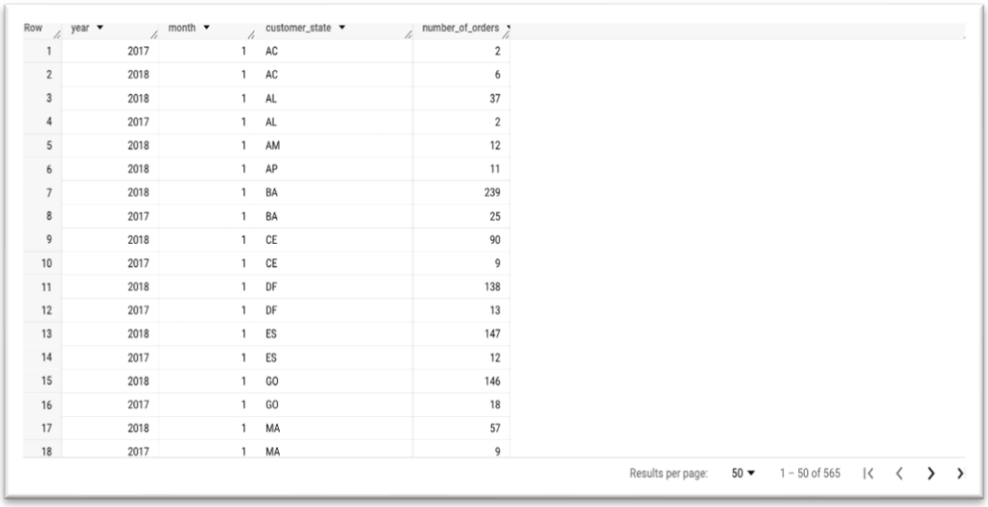
III. Evolution of E-commerce orders in the Brazil region:

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

QUERY:

```
SELECT
EXTRACT (YEAR FROM B.order_purchase_timestamp) AS year,
EXTRACT (MONTH FROM B.order_purchase_timestamp) AS month,
A.customer_state,
COUNT (*) AS number_of_orders
FROM
`E_COMMERCE_TARGET.Customers` A
join `E_COMMERCE_TARGET.Orders` B
on A.customer_id =B.customer_id
GROUP BY
year,
month,
A.customer_state
ORDER BY
month,
A.customer_state;
```

OUTPUT:



Row	year	month	customer_state	number_of_orders
1	2017	1	AC	2
2	2018	1	AC	6
3	2018	1	AL	37
4	2017	1	AL	2
5	2018	1	AM	12
6	2018	1	AP	11
7	2018	1	BA	239
8	2017	1	BA	25
9	2018	1	CE	90
10	2017	1	CE	9
11	2018	1	DF	138
12	2017	1	DF	13
13	2018	1	ES	147
14	2017	1	ES	12
15	2018	1	GO	146
16	2017	1	GO	18
17	2018	1	MA	57
18	2017	1	MA	9

Results per page: 50 1 – 50 of 565

INSIGHTS:

The above table shows the varying order counts across different months in different states, indicating potential seasonal trends or fluctuations in customer demand throughout the year. The top state getting consistent higher orders is "SP" and the bottom states which are getting lower orders are 'RR', 'AP', 'AM' and so on.

RECOMMENDATIONS:

- By understanding the regional differences from the above table, 'Target' can undertake state-specific marketing campaigns that resonate with local preferences and needs.
- The other reason could be Impact of festive seasons in certain months i.e., seasonal spikes in certain months, such as around holidays or festivals, may result in increased order activity. 'Target' can capitalize on these peak periods with relevant promotions and incentives.

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

QUERY:

```
SELECT
customer_state,
Count (*) as no_of_customers
from `E_COMMERCE_TARGET.Customers`
group by customer_state
order by no_of_customers desc;
```

OUTPUT:

Row	customer_state	no_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
11	PE	1652
12	CE	1336
13	PA	975
14	MT	907
15	MA	747
16	MS	715
17	PB	536
18	PI	495

Results per page: 50 1 – 27 of 27

INSIGHTS:

The above table provides the count of customers for each state from 'Customers' table. SP, RJ and MG states have the larger customer base and represent significant market opportunities. RR, AP, AC states have the lower customer count and represent lower customer base and market activities

RECOMMENDATIONS:

- Expansion of Opportunities in lower count customer states and follow targeted marketing and advertising strategies, localized strategies help in understanding the local demand of customers.
- **Targeted Marketing:** Focus marketing efforts on states with a higher count of customers to capitalize on the existing customer base and drive more sales. Compare the customer count across states with the performance of competitors in those regions to gain insights into market share and identify potential areas for improvement.

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

A. 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:

```
with cte as (
select
Extract (year from order_purchase_timestamp) as year,
Extract (month from order_purchase_timestamp) as month,
Round (sum(p.payment_value),2) as cost_of_orders
from `E_COMMERCE_TARGET.Payments` p
join `E_COMMERCE_TARGET.Orders` o
on p.order_id = o.order_id
group by year, month
having year in (2017,2018) and month between 1 and 8
order by month),

cte2 as (
select cte.year,
Sum (cost_of_orders) as cost_order
from cte
group by cte.year
order by cte.year)

select cte2.year, cte2.cost_order,
Round ((((cost_order - lag(cost_order,1) over(order by cte2.year))/lag(cost_order,1) over(order by cte2.year))),2) * 100 as percentage_increase
from cte2;
```

OUTPUT:

Row	year	cost_order	percentage_increase
1	2018	8694733.84	137.0
2	2017	3669022.12	null

ASSUMPTIONS:

`Cost_order` represents the cost of total orders in the year.

INSIGHTS:

If we observe the above table in 2017, we got cost of orders = 3669022 and in the year 2018 we got cost of orders = 8694733 which is higher compared to the previous year. So, the percentage increase that we get for 2017-2018 is 136.98 (~ 137).

RECOMMENDATIONS:

Ensure that the supply chain is optimized to handle the higher demand growth efficiently, preventing stockouts and ensuring timely deliveries. Seasonal marketing campaigns tailored to customer spending behavior during different months of the year.

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

QUERY:

```
select A.customer_state,
Round (sum(C.price),2) as State_sum_price,
Round (Avg(C.price),2) as State_avg_price
from `E_COMMERCE_TARGET.Customers` A
join `E_COMMERCE_TARGET.Orders` B
on A.customer_id = B.customer_id
join `E_COMMERCE_TARGET.Order_items` C
on C.order_id = B.order_id
group by A.customer_state
order by State_sum_price asc,State_avg_price asc;
```

OUTPUT:

Row	customer_state	State_sum_price	State_avg_price
1	RR	7829.43	150.57
2	AP	13474.3	164.32
3	AC	15982.95	173.73
4	AM	22356.84	135.5
5	RO	46140.64	165.97
6	TO	49621.74	157.53
7	SE	58920.85	153.04
8	AL	80314.81	180.89
9	RN	83034.98	156.97
10	PI	86914.08	160.36
11	PB	115268.08	191.48
12	MS	116812.64	142.63
13	MA	119648.22	145.2
14	MT	156453.53	148.3
15	PA	178947.81	165.69
16	CE	227254.71	153.76
17	PE	262788.03	145.51
18	ES	275037.31	121.91

Results per page: 50 1 – 27 of 27

ASSUMPTUIONS:

- `State_sum_price` represents the total order price associated with each state.
- `State_avg_price` represents the average order price associated with each state.

INSIGHTS:

- The above table provides insights into the total price (state_sum_price) and average price (state_avg_price) of orders placed by customers in each state. “States with higher State sum price have customers making more significant purchases, while states with higher State avg price indicate customers with a tendency to spend more per order”.
- ‘SP’ state has the highest state_sum_price (5202955.05) and ‘RR’ state has the lowest state_sum_price (7829.43) of all states. ‘PB’ state has the highest state_avg_price (191.48) and ‘SP’ state has the lowest state_avg_price (109.65).

RECOMMENDATIONS:

- Use the price insights to determine which products or price ranges are more appealing to customers in specific states. Customize product offerings to suit the preferences of each region.
- Segment customers based on their states and price behavior to tailor communication and promotions to different customer groups. Identify states with higher State avg prices to

explore upselling opportunities by recommending complementary or premium products to increase the order value.

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

QUERY:

```
select A.customer_state,
Round (sum(C.freight_value),2) as State_freight_sum,
Round (Avg(C.freight_value),2) as State_freight_avg
from `E_COMMERCE_TARGET.Customers` A
join `E_COMMERCE_TARGET.Orders` B
on A.customer_id = B.customer_id
join `E_COMMERCE_TARGET.Order_items` C
on C.order_id = B.order_id
group by A.customer_state
order by State_freight_sum asc,State_freight_avg;
```

OUTPUT:

Row	customer_state	State_freight_sum	State_freight_avg
1	RR	2235.19	42.98
2	AP	2788.5	34.01
3	AC	3686.75	40.07
4	AM	5478.89	33.21
5	RO	11417.38	41.07
6	TO	11732.68	37.25
7	SE	14111.47	36.65
8	AL	15914.59	35.84
9	RN	18860.1	35.65
10	MS	19144.03	23.37
11	PI	21218.2	39.15
12	PB	25719.73	42.72
13	MT	29715.43	28.17
14	MA	31523.77	38.26
15	PA	38699.3	35.83
16	CE	48351.59	32.71
17	ES	49764.6	22.06
18	DF	50625.5	21.04

Results per page: 50 1 – 27 of 27

ASSUMPTIONS:

- `State_freight_sum` represents the total amount of freight within a specific state.
- `State_freight_avg` represents the average amount of freight within a specific state.

INSIGHTS:

States with higher State_freight_sum have customers incurring more significant shipping costs, while states with higher State_freight_avg indicate customers with relatively higher shipping and trasportation charges. From the above table we can observe that ‘SP’ state has highest

State_freight_sum (718723.07) and 'RR' state has lowest State_freight_sum (2235.19). Similarly the states with highest and lowest State_freight_avg are 'SP' (15.15) and 'RR' (42.98) respectively.

RECOMMENDATIONS:

For states with higher freight costs, businesses should focus on improving customer service to justify shipping expenses and enhance customer satisfaction and leverage the insights to negotiate bulk shipping discounts with logistics partners, especially for states with higher State_freight_avg, to reduce the average shipping cost per order.

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

A. Find the no. of days taken to deliver each order from the order's purchase date as delivery time.
Also, calculate the difference (in days) between the estimated & actual delivery date of an order.
Do this in a single query.

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,
timestamp_diff (order_delivered_customer_date, order_purchase_timestamp, day) as
time_to_deliver_in_days,
timestamp_diff (order_estimated_delivery_date, order_delivered_customer_date, day) as
Actual_and_estimated_delivery_diff
from `E_COMMERCE_TARGET.Orders`
order by order_purchase_timestamp;
```

OUTPUT:

Row	order_id	time_to_deliver_in_days	Actual_and_estimated_delivery_diff
1	0607f0efea4b566f1eb8f7d3c2...	3	146
2	c72727d29cde4cf870d569bf6...	6	139
3	eec7f369423b033e549c02f3c...	20	134
4	c2bb89b5c1dd978d507284be7...	16	123
5	40dc2ba6f322a17626aac6244...	7	108
6	1a695d543b7302aa9446c8d5f...	12	83
7	39e0115911bf404857e14baa7...	11	82
8	38930f76efb00b138f4d632e4d...	11	77
9	c5132855100a12d63ed4e8ae0...	12	77
10	559eea5a72341a4c82dbce988...	13	77
11	cfe2129968372224ea392ad47...	14	76
12	e2ea01f68f92f0c55b08ca739a...	19	75
13	51fa5b2b3ea1aa99dc3cea8f00...	8	74
14	39ff1e358efa92ceae42de6d5f...	11	74
15	7b875e17a8b649ecb72124320...	18	72
16	e03aeb40ff9fb4a40808ce5083...	6	72
17	440b5d8129165b7e04b3c37e3...	22	70
18	538cce8b584cc6b13e623e781...	12	70
19	35fbd0c1666a2c04e7d111da0...	6	68
20	09c636e254bb18eeef9434520...	28	68
21	da44f554b7c4550b0d3cd650f...	15	68
22	c43c85be1846e97fe6554a688...	18	67

Results per page: 50 1 – 50 of 99441

ASSUMPTIONS:

- The `time_to_deliver_in_days` column represents the number of days it took for the order to be delivered to the customer from the purchase date.
- The `Actual_and_estimated_delivery_diff` column measures the difference in days between the estimated delivery date and the actual delivery date.

INSIGHTS:

This information can help us investigate the reasons behind these delays and take appropriate actions to improve delivery performance.

RECOMMENDATIONS:

Regularly monitor delivery metrics, track performance over time, and make data-driven decisions to enhance the overall delivery experience. Continuously seek customer feedback and leverage it to make improvements.

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

QUERY:

```
select A.Top_states,
A.Top5_avg_freight_value,
A.Top_5,
B.Bottom_states,
B.Bottom5_avg_freight_value,
B.Bottom_5
from
(select *
from
(select customer_state as Top_states,
Round (avg(freight_value),2) as Top5_avg_freight_value,
dense_rank() over(order by round(avg(freight_value),2)desc) as Top_5
from `E_COMMERCE_TARGET.Orders` o
join `E_COMMERCE_TARGET.Customers` c
on o.customer_id = c.customer_id
join `E_COMMERCE_TARGET.Order_items` oi
on o.order_id = oi.order_id
group by customer_state)
where Top_5 <= 5) A

join

(select *
from
(select customer_state as Bottom_states,
Round (avg(freight_value),2) as Bottom5_avg_freight_value,
dense_rank() over(order by round(avg(freight_value),2)) as Bottom_5
from `E_COMMERCE_TARGET.Orders` o
join `E_COMMERCE_TARGET.Customers` c
on o.customer_id = c.customer_id
join `E_COMMERCE_TARGET.Order_items` oi
on o.order_id = oi.order_id
group by customer_state)tbl
where Bottom_5 <= 5) B

on A.Top_5 = B.Bottom_5
order by Top_5;
```

OUTPUT:

Row	Top_states	Top5_avg_freight_value	Top_5	Bottom_states	Bottom5_avg_freight_value	Bottom_5
1	RR	42.98	1	SP	15.15	1
2	PB	42.72	2	PR	20.53	2
3	RO	41.07	3	MG	20.63	3
4	AC	40.07	4	RJ	20.96	4
5	PI	39.15	5	DF	21.04	5

ASSUMPTIONS:

- `Top5_avg_freight_value` represents the average value of freight or cargo shipments for the top 5 states with the highest freight values.
- `Bottom5_avg_freight_value` represents the average value of freight or cargo shipments for the bottom 5 states with the lowest freight value.

INSIGHTS:

The above table provides insights of top 5 and bottom 5 states performance in average freight value. The top 5 states (RR, PB, RO, AC, PI,) have a higher average freight value and bottom 5 states (SP, PR, MG, RJ, DF) have lower average freight value. This information can help us understand potential challenges in logistics and transportation costs.

RECOMMENDATIONS:

Based on the insights, focus on optimizing freight costs in states with high average freight values. Explore options like negotiating bulk shipping discounts with logistics partners, especially for states with higher freight_avg, optimizing packaging, or leveraging local distribution centers to reduce the average shipping cost per order.

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

QUERY:

```
select A.Top5_states,
A.Top5_avg_delivery_time_in_days,
A.Top_5,
B.Bottom5_states,
B.Bottom5_avg_delivery_time_in_days,
```

```
B.Bottom_5
from
(select *
from
(select customer_state as Top5_states,
Round (avg(timestamp_diff(order_delivered_customer_date, order_purchase_timestamp, day)),2) as Top5_avg_delivery_time_in_days,
dense_rank() over(order by round(avg(timestamp_diff(order_delivered_customer_date, order_purchase_timestamp, day)),2) desc) as Top_5
from `E_COMMERCE_TARGET.Orders` o
join `E_COMMERCE_TARGET.Customers` c
on o.customer_id = c.customer_id
join `E_COMMERCE_TARGET.Order_items` oi
on o.order_id = oi.order_id
group by customer_state
order by Top5_avg_delivery_time_in_days desc)
where Top_5 <= 5) A

join

(select *
from
(select customer_state as Bottom5_states,
Round (avg(timestamp_diff(order_delivered_customer_date, order_purchase_timestamp, day)),2) as Bottom5_avg_delivery_time_in_days,
dense_rank() over(order by round(avg(timestamp_diff(order_delivered_customer_date, order_purchase_timestamp, day)),2)) as Bottom_5
from `E_COMMERCE_TARGET.Orders` o
join `E_COMMERCE_TARGET.Customers` c
on o.customer_id = c.customer_id
join `E_COMMERCE_TARGET.Order_items` oi
on o.order_id = oi.order_id
group by customer_state
order by Bottom5_avg_delivery_time_in_days)
where Bottom_5<= 5) B

on A.Top_5=B.Bottom_5
order by A.Top_5;
```

OUTPUT:

Row	Top5_states	Top5_avg_delivery_time_in_days	Top_5	Bottom5_states	Bottom5_avg_delivery_time_in_days	Bottom_5
1	RR	27.83	1	SP	8.26	1
2	AP	27.75	2	PR	11.48	2
3	AM	25.96	3	MG	11.52	3
4	AL	23.99	4	DF	12.5	4
5	PA	23.3	5	SC	14.52	5

ASSUMPTIONS:

- `Top5_avg_delivery_time_in_days` represents the average time it takes for deliveries to be completed for the top 5 states with the fastest delivery times.
- `Bottom5_avg_delivery_time_in_days` represents the average time it takes for deliveries to be completed for the bottom 5 states with the slowest delivery times.

INSIGHTS:

By comparing the average delivery times for the top 5 states and bottom 5 states, we can identify that the states RR, AP, AM, AL, PA are slow delivering states and the states SP, PR, MG, DF, SC are fast delivering states. This information can help us understand potential challenges in the logistics and delivery process.

RECOMMENDATIONS:

- Examine states with longer delivery times (B. Bottom5_states) to pinpoint potential bottlenecks or issues affecting the delivery process. Analyze the data of the top 5 states and try to implement their practices to improve the delivery time of bottom 5 states.
- Encouraging customers to provide regular feedback on their delivery experiences and addressing the challenges related to delivery processes can improve overall customer satisfaction.

D. 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:

```
with top5 as
(
select B.customer_state as state,A.order_status,
Round (avg(timestamp_diff(order_estimated_delivery_date, order_delivered_customer_date, day)),2) as time_diff_in_days
from `E_COMMERCE_TARGET.Orders` A
join `E_COMMERCE_TARGET.Customers` B
on A.customer_id = B.customer_id
group by B.customer_state,A.order_status
having A.order_status = 'delivered'
```

```
order by time_diff_in_days desc
)
select *
from (select state,
time_diff_in_days,
dense_rank() over(order by top5.time_diff_in_days desc) as Fast_delivering_states_rank
from top5) F
where F.Fast_delivering_states_rank <= 5
order by Fast_delivering_states_rank;
```

OUTPUT:

Row	state	time_diff_in_days	Fast_delivering_states_rank
1	AC	19.76	1
2	RO	19.13	2
3	AP	18.73	3
4	AM	18.61	4
5	RR	16.41	5

ASSUMPTIONS:

- `time_diff_in_days` represents the difference in time, measured in days, between order_estimated_delivery_date and order_delivered_customer_date.
- `Fast_delivering_states_rank' represents ranking of states based on their speed of delivery.

INSIGHTS:

The above table provides insights of the states that have the fastest delivery time for delivered orders.

The top 5 states that have the fastest delivery time are:

AC = 19.76 days

RO = 19.13 days

AP = 18.73 days

AM = 18.61 days

RR = 16.41 days

RECOMMENDATIONS:

Focus on improving last-mile delivery logistics in states that rank lower in fast delivering states. Collaborate with local delivery partners to enhance the efficiency of the final stages of the delivery process. Consider implementing performance-based incentives for delivery partners

and staff in regions with the fastest delivery times. Encourage and reward exceptional delivery performance to maintain high delivery standards.

VI. Analysis based on the payments:

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

QUERY:

```
SELECT
EXTRACT (MONTH FROM A.order_purchase_timestamp) AS month,
COUNT (*) AS number_of_orders_Count,
B.payment_type
FROM `E_COMMERCE_TARGET.Orders` A
join `E_COMMERCE_TARGET.Payments` B
on A.order_id =B.order_id
GROUP BY
month,
payment_type
ORDER BY
month;
```

OUTPUT:

Row	month	number_of_orders_Count	payment_type
1	5	8350	credit_card
2	8	8269	credit_card
3	7	7841	credit_card
4	3	7707	credit_card
5	4	7301	credit_card
6	6	7276	credit_card
7	2	6609	credit_card
8	1	6103	credit_card
9	11	5897	credit_card
10	12	4378	credit_card
11	10	3778	credit_card
12	9	3286	credit_card
13	8	2077	UPI
14	7	2074	UPI
15	5	2035	UPI
16	3	1942	UPI
17	6	1807	UPI
18	4	1783	UPI

Results per page: 50 1 – 50 of 50

INSIGHTS:

By observing the above table, we can identify that credit card payment type is the highest preferred mode of transaction in 5th month, 2nd comes UPI in 8th month, 3rd comes voucher in 7th month and last preferred is debit card in 8th month, subsequently there are 2 other undefined mode of transactions.

RECOMMENDATIONS:

Offer a variety of secure and convenient payment options based on the popularity of different payment types. Ensure that your checkout process is user-friendly and supports the preferred payment methods of your customers.

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

QUERY:

```
SELECT
COUNT (*) AS number_of_orders_Count,
B.payment_installments
FROM `E_COMMERCE_TARGET.Orders` A
join `E_COMMERCE_TARGET.Payments` B
on A.order_id =B.order_id
GROUP BY
payment_installments
having B.payment_installments <> 0
ORDER BY
payment_installments;
```

OUTPUT:

Row	number_of_orders_Count	payment_installments
1	52546	1
2	12413	2
3	10461	3
4	7098	4
5	5239	5
6	3920	6
7	1626	7
8	4268	8
9	644	9
10	5328	10
11	23	11
12	133	12
13	16	13
14	15	14
15	74	15
16	5	16
17	8	17
18	27	18

Results per page: 50 1 – 23 of 23

INSIGHTS:

The above table provides insights regarding the payment installments chosen by the customers. From the above table the highest number of customers chose to pay in 1, 2 and 3 installments.

RECOMMENDATIONS:

Based on the popularity of different installment options, consider offering flexible installment plans tailored to customer needs. Allow customers to choose from various installment durations and interest rates, if applicable.

SUBMITTED BY

B. RAJU NAIK