TARGET SQL BUSINESS CASE

By- Avanti Raut

Q1.Import the dataset and do usual exploratory analysis steps like checking the structure & characteristics of the dataset:

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

Ans:

There are total 5 columns in “customers” table -

Column Name Data type

1. customer_id String

2. customer_unique_id String

3. customer_zip_code_prefix INT64

4. customer_city String

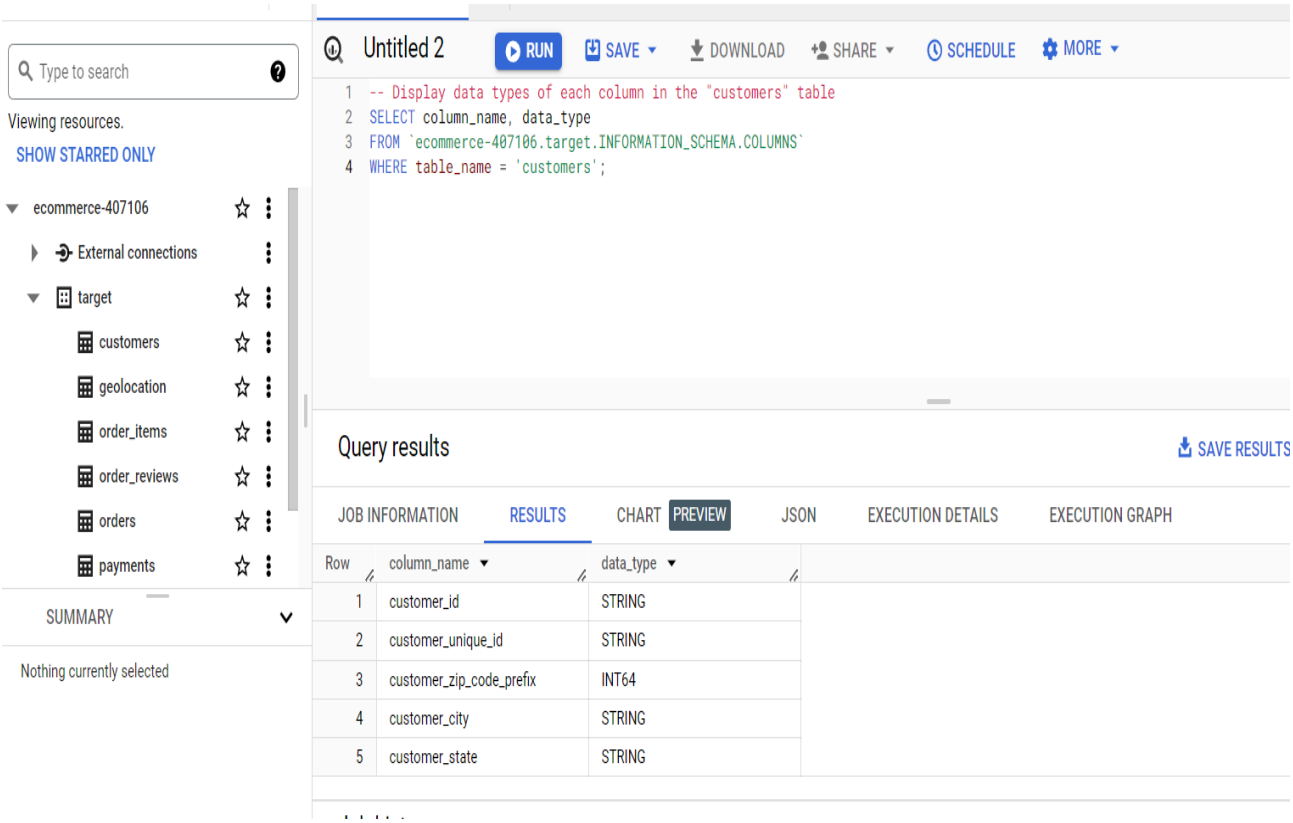
5. customer_state String

Here is the query

```
-- Display data types of each column in the "customers" table
SELECT column_name, data_type
FROM `ecommerce-407106.target.INFORMATION_SCHEMA.COLUMNS`
WHERE table_name = 'customers';
```

HERE is the screenshot of the query to see the data type in bigquery for

“customers” table



Insights: there are a total of 5 columns in this table. 4 column are string and on is INT64

Recommendation: NA

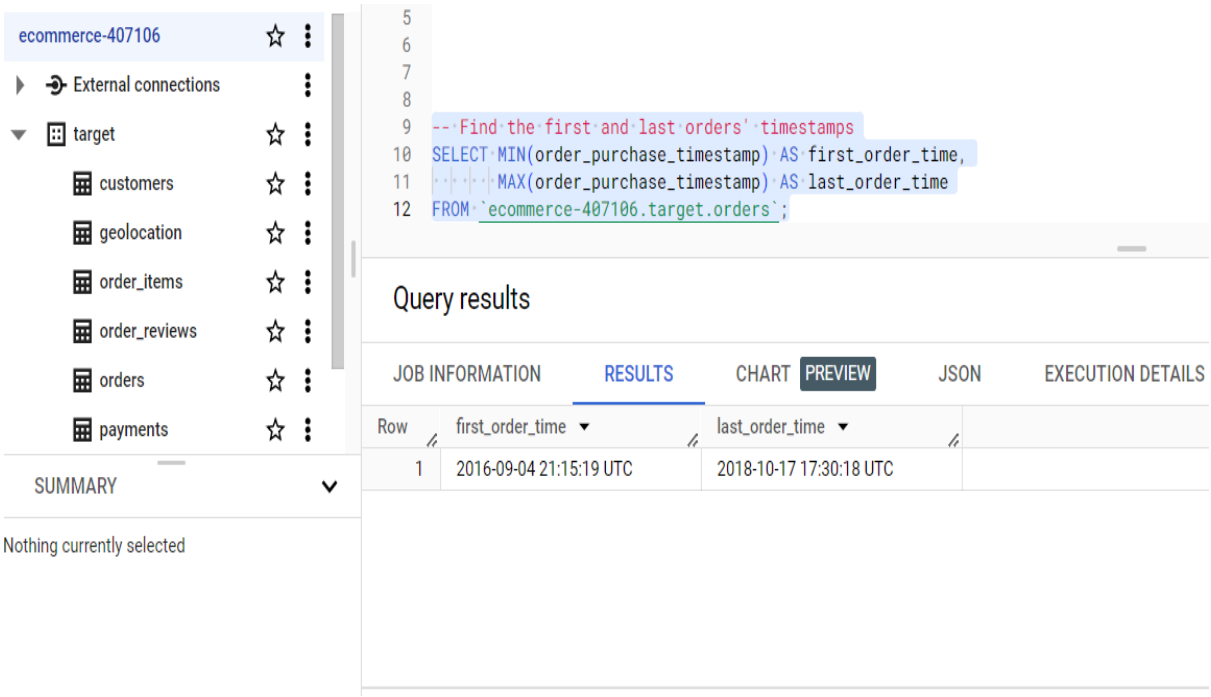
1.(II).Get the time range between which the orders were placed.

Ans:

Here is the query

```
-- Find the first and last orders' timestamps
SELECT MIN(order_purchase_timestamp) AS first_order_time,
       MAX(order_purchase_timestamp) AS last_order_time
FROM `ecommerce-407106.target.orders`;
```

Here I am adding the screenshot of the query:



Insights: First order was placed on “2016-09-04 21:15:19 UTC” and last order was placed on “2018-10-17 17:30:18 UTC” according to the “target” dataset.

Recommendation: NA

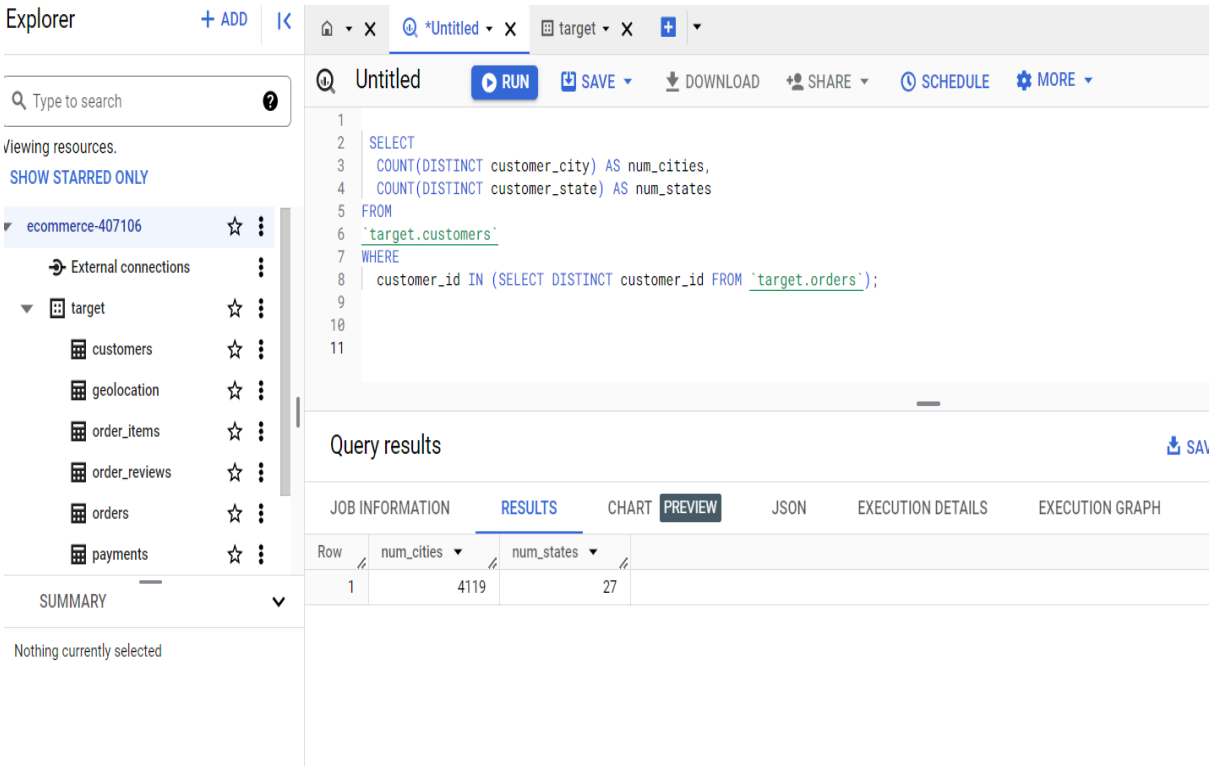
1.(III).Count the Cities & States of customers who ordered during the given period.

Ans:

Here is the query:

```
SELECT
  COUNT(DISTINCT customer_city) AS num_cities,
  COUNT(DISTINCT customer_state) AS num_states
FROM
`target.customers`
WHERE
  customer_id IN (SELECT DISTINCT customer_id FROM `target.orders`);
```

Here is the screenshot of query result



Insights: During the given time period there are 4119 and 27 states from where orders were placed.

Recommendation: NA

Q2. In-depth Exploration:

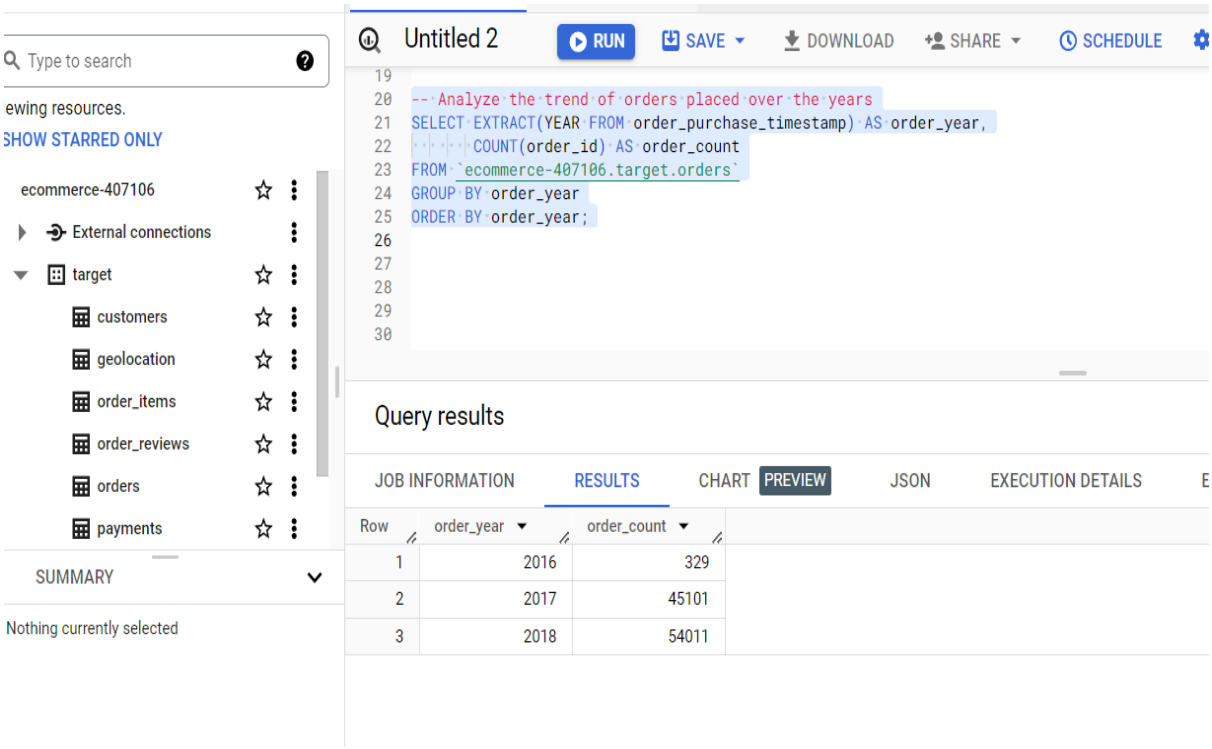
2.(I). Is there a growing trend in the no. of orders placed over the past years?

Ans:

HERE GOES THE QUERY:

```
-- Analyze the trend of orders placed over the years
SELECT EXTRACT(YEAR FROM order_purchase_timestamp) AS order_year,
       COUNT(order_id) AS order_count
FROM `ecommerce-407106.target.orders`
GROUP BY order_year
ORDER BY order_year;
```

And here is the screenshot of the result



Insights: yes it is clearly visible that there is continuous growth in the number of orders.

In 2016 there were only 329 orders but the next year(2017) it went up to 45001.

In 2018 this number again got a huge hike and a total of 54011 orders were placed.

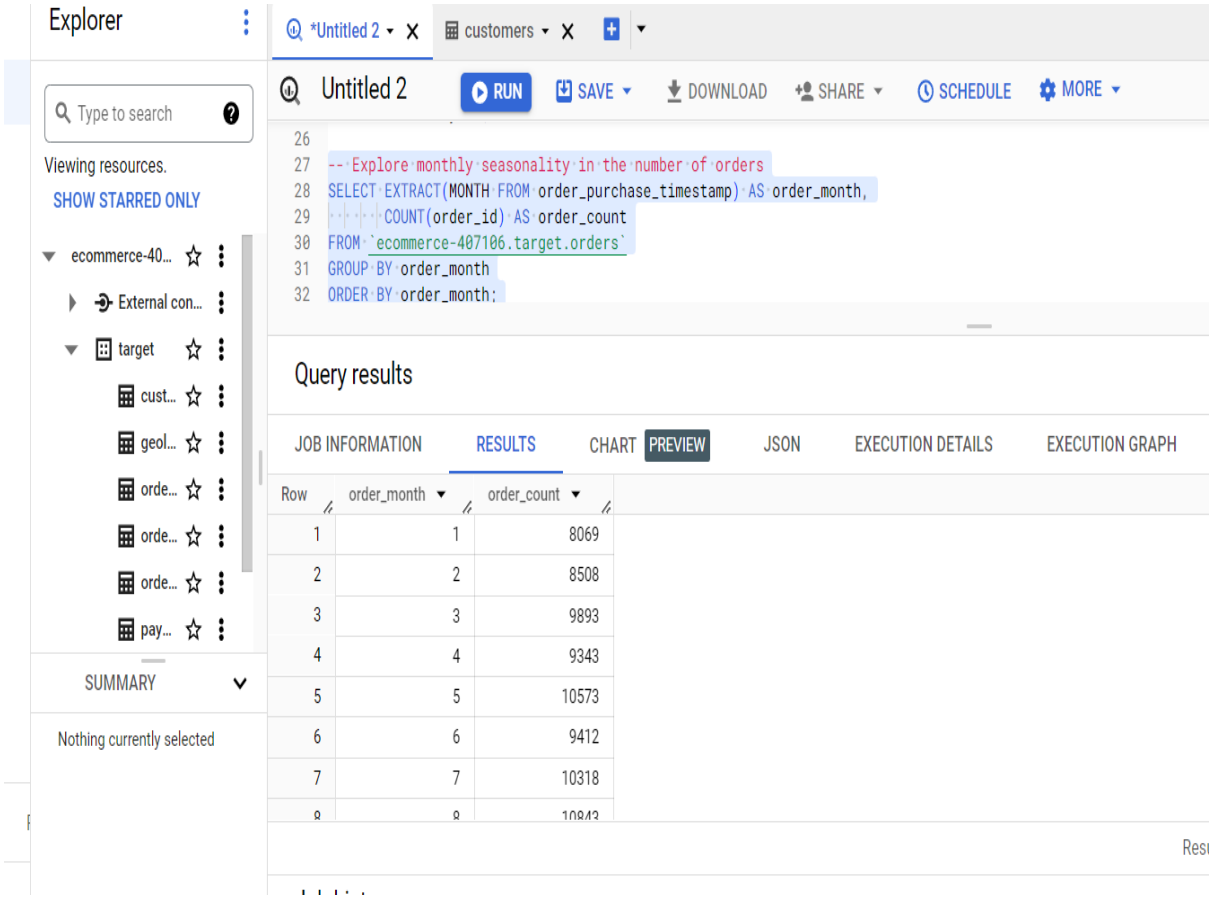
Recommendation: As we can see a trend where no. of orders are increasing year by year so “target” should also focus on the other aspects like delivery time and customer service so customers can be retained.

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

Here goes the query:

```
-- Explore monthly seasonality in the number of orders
SELECT EXTRACT(MONTH FROM order_purchase_timestamp) AS order_month,
       COUNT(order_id) AS order_count
FROM `ecommerce-407106.target.orders`
GROUP BY order_month
ORDER BY order_month;
```

Here is the screenshot of results



Insights: If we compare month on month data for total order during 2016 to 2018.

We see a trend that order number increases from January to March gradually then it takes a little dip in April but again it goes high in march and crosses 10k orders.

During this whole period “Target” got the highest number of orders in the month of August (10843) orders.

Recommendation: As there is not a big pattern in monthly data so we can also look into Quarterly trend then we can say Q2 got the highest orders(29328) then comes Q1 (26470) orders followed by Q3 (25466) orders then Q4 comes last with (18177) orders.

So here we can see that there are less orders in Q3 AND Q4 so we can make our marketing budget in these Quarters next year or we should plan something like better deals and discounts during this period so we can get the same number of orders.

Q2(III).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

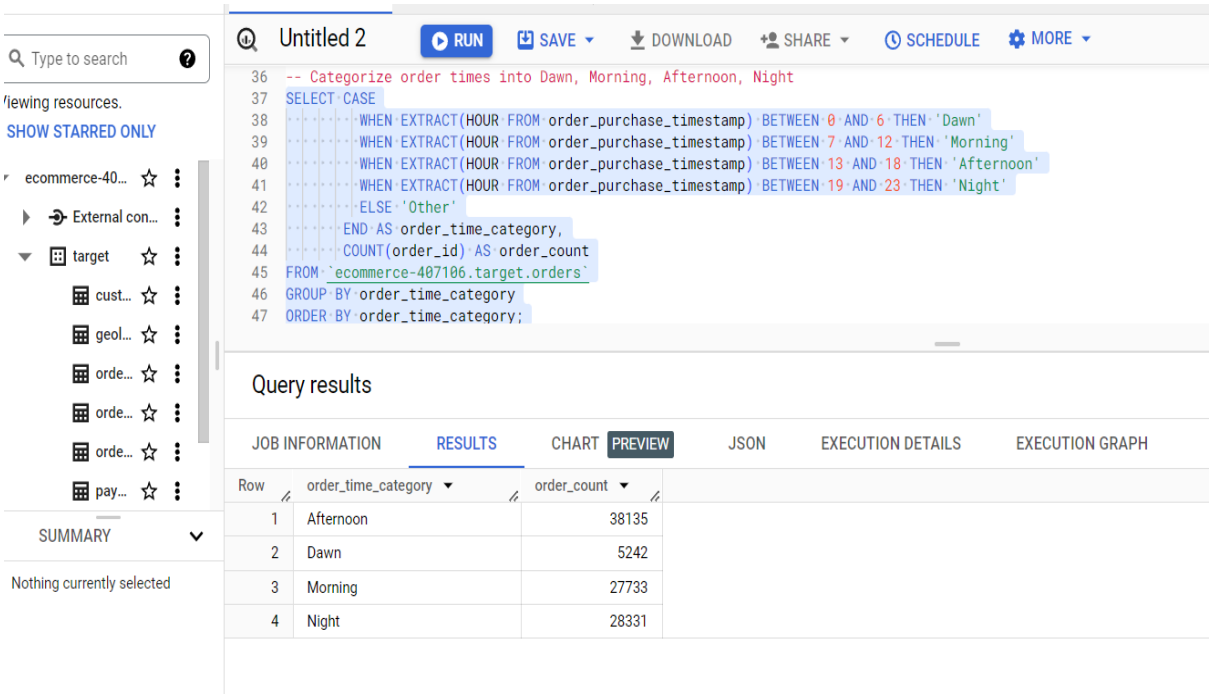
● 19-23 hrs : Night

Ans: As per the analysis customers placed most number of orders in “Afternoon” (38135) then 2nd highest order placed in “Night” (28331) followed by “Mornings” (27733) orders then least number of orders placed during “Dawn” (5242).

Here goes the 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'
         ELSE 'Other'
       END AS order_time_category,
       COUNT(order_id) AS order_count
FROM `ecommerce-407106.target.orders`
GROUP BY order_time_category
ORDER BY order_time_category;
```

Here is the screenshot of results:



Insights: As per the analysis customers placed most number of orders in “Afternoon” (38135) then 2nd highest order placed in “Night” (28331) followed by “Mornings” (27733) orders then least number of orders placed during “Dawn” (5242).

Recommendation: As we can see that customers mostly placed orders in afternoon and mornings so we can focus more on these times in terms of ad spend and marketing.

Q3.Evolution of E-commerce orders in the Brazil region:

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

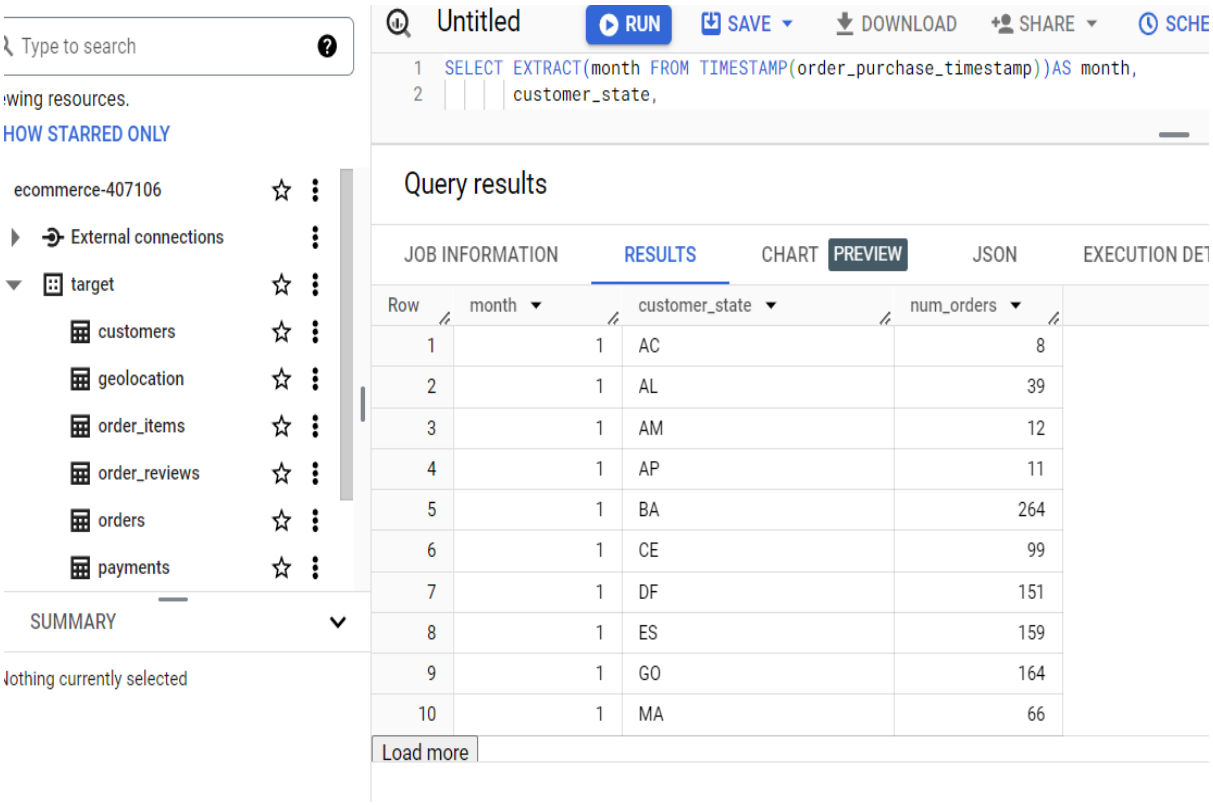
Ans:

Here goes the query:

```
SELECT EXTRACT(month FROM TIMESTAMP(order_purchase_timestamp))AS month,
       customer_state,
       count(*) As num_orders
FROM `target.orders` o
JOIN `target.customers` c
on o.customer_id=c.customer_id
Group by month,customer_state
```

```
order by month,customer_state
```

Here is the screenshot of result:



Insights: We got the month on month data by extracting month from timestamp then doing left join gives all the orders details.

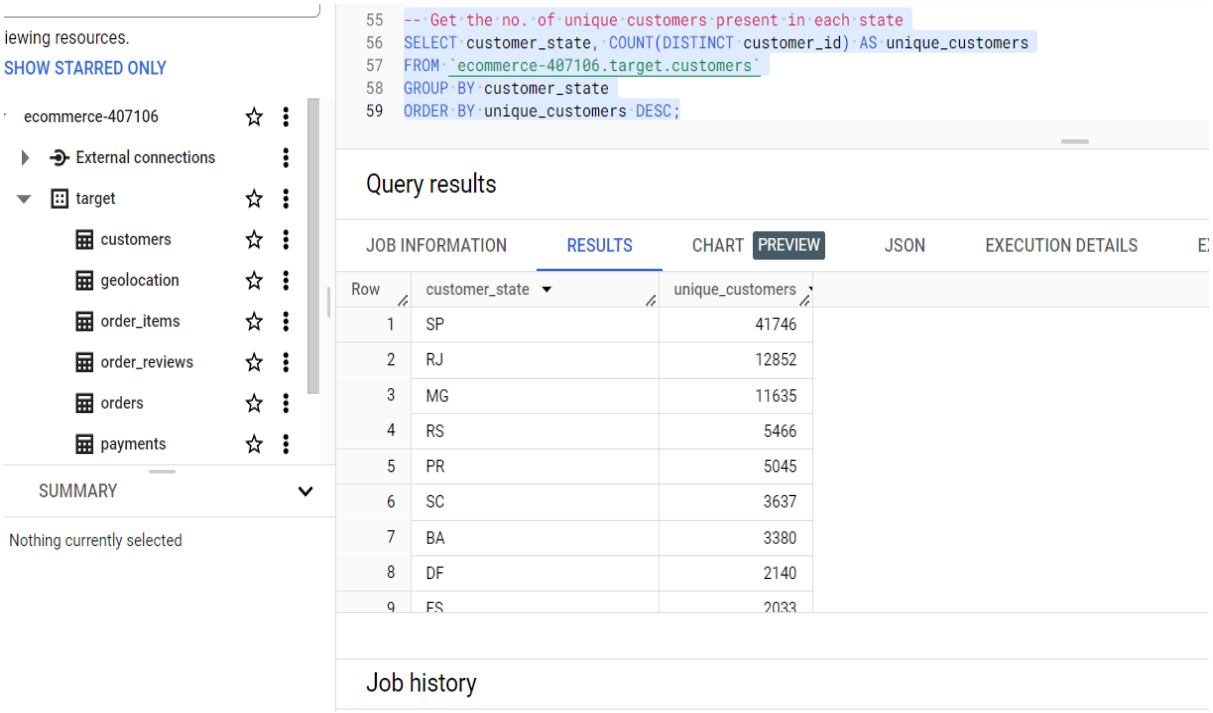
Recommendations: NA

3.(II).How are the customers distributed across all the states?

Here goes the query:

```
-- Get the no. of unique customers present in each state
SELECT customer_state, COUNT(DISTINCT customer_id) AS unique_customers
FROM `ecommerce-407106.target.customers`
GROUP BY customer_state
ORDER BY unique_customers DESC;
```

Here is the screenshot of results:



Insights: To get this data we can count distinct customer_id and customer_state from customers table. Then grouping by state will give the exact result.

Recommendations: NA

Q4.Impact on Economy: Analyze the money movement by e-commerce by looking at order prices, freight and others.

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

Ans: Cost of order in 2017 between January to August was “3669022.12” and during the same period cost in 2018 was “8694733.84”

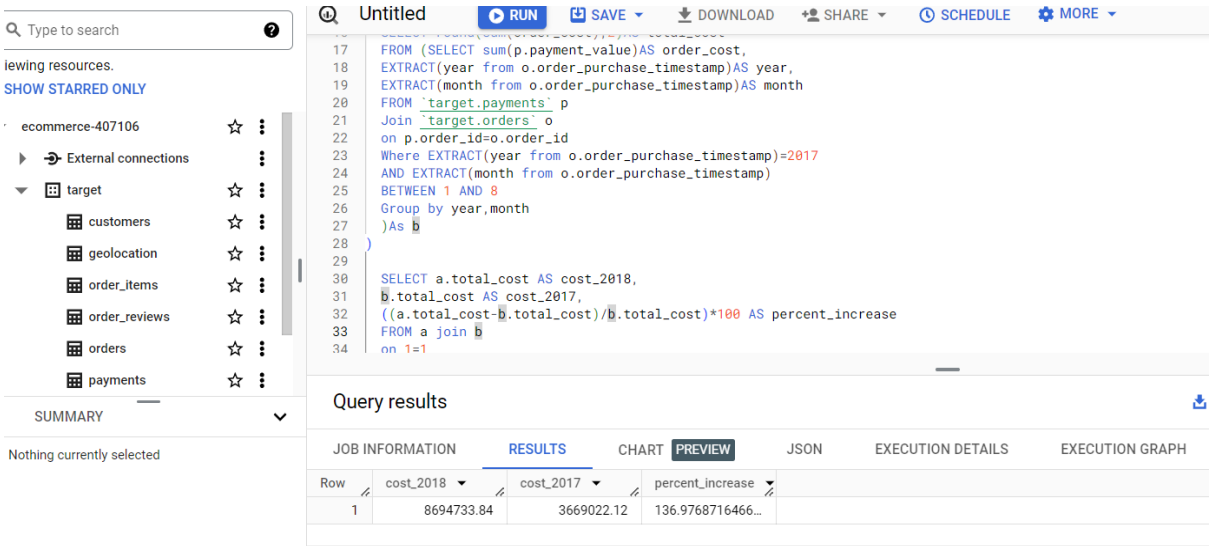
So if we calculate increased % then it is 136.97%

Here is the query :

```sql
With a as(
  SELECT round(sum(order_cost),2)AS total_cost
  FROM (SELECT sum(p.payment_value)AS order_cost,
  EXTRACT(year from o.order_purchase_timestamp)AS year,
  EXTRACT(month from o.order_purchase_timestamp)AS month
  FROM `target.payments` p
  Join `target.orders` o
  on p.order_id=o.order_id
  WHERE EXTRACT(year from o.order_purchase_timestamp)=2018
  AND EXTRACT(month from o.order_purchase_timestamp)
  BETWEEN 1 AND 8
  Group by year,month
  )AS a
),
b as(
  SELECT round(sum(order_cost),2)AS total_cost
  FROM (SELECT sum(p.payment_value)AS order_cost,
  EXTRACT(year from o.order_purchase_timestamp)AS year,
  EXTRACT(month from o.order_purchase_timestamp)AS month
  FROM `target.payments` p
  Join `target.orders` o
  on p.order_id=o.order_id
  Where EXTRACT(year from o.order_purchase_timestamp)=2017
  AND EXTRACT(month from o.order_purchase_timestamp)
  BETWEEN 1 AND 8
  Group by year,month
  )As b
)

  SELECT a.total_cost AS cost_2018,
  b.total_cost AS cost_2017,
  ((a.total_cost-b.total_cost)/b.total_cost)*100 AS percent_increase
  FROM a join b
  on 1=1
```

Here is the screenshot of result:



Insights: Cost of order in 2017 between January to August was “3669022.12” and during the same period cost in 2018 was “8694733.84”

So if we calculate increased % then it is 136.97%

Recommendations: The substantial increase of 136.97% in the cost of orders from 2017 to 2018 indicates positive business growth. This growth may be attributed to increased customer demand, expanded product offerings, or successful marketing strategies.

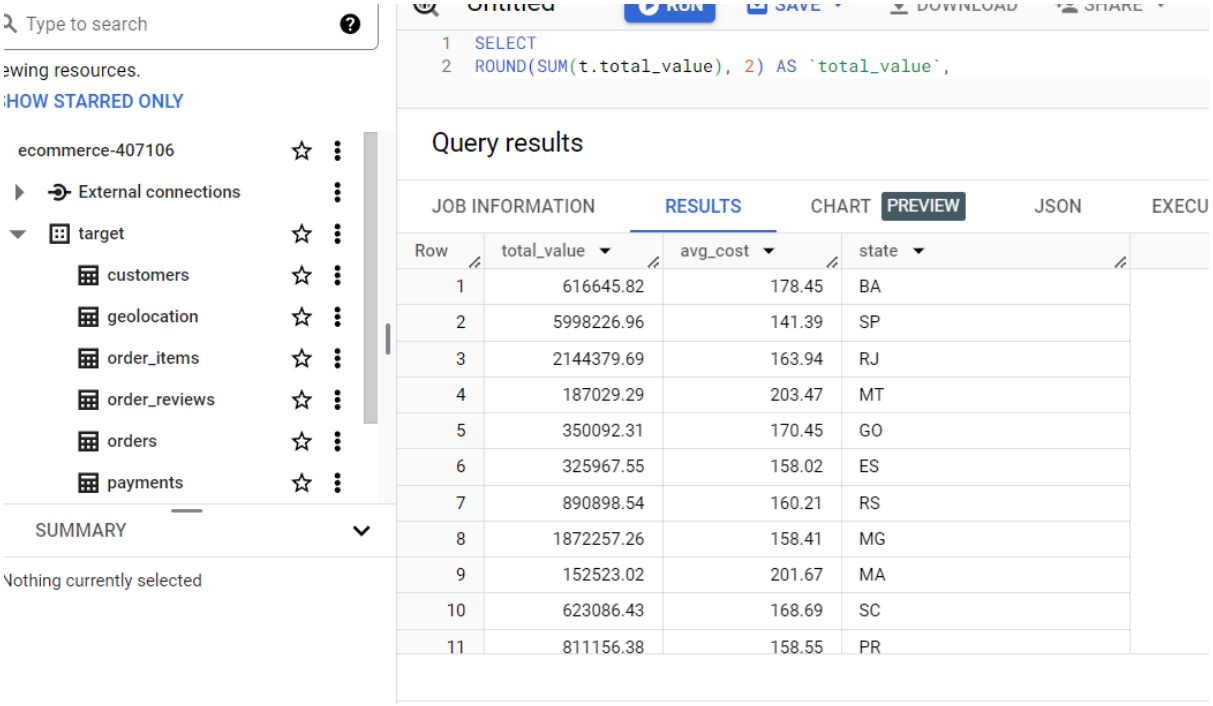
With increased business, it's also crucial to focus on operational efficiency. Evaluate whether current processes can handle the growth effectively or if there are areas for improvement to streamline operations.

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

Here is the query:

```
SELECT round(sum(t.total_value),2)AS `total_value`,
       round(avg(t.avg_cost),2)AS `avg_cost`,
       t.state
FROM (SELECT distinct(o.order_id),
round(sum(p.payment_value),2)AS `total_value`,
round(avg(p.payment_value),2)AS `avg_cost`,
c.customer_state AS `state`
FROM `target.payments` p
join `target.orders` o
on p.order_id=o.order_id
join `target.customers` c
on o.customer_id=c.customer_id
Group by order_id,`state`
)t
Group by t.state;
```

Here is the screenshot of result:



Insights: to calculate the total value we can sum the payment_value column.

And to find avg cost we can take the avg of payment_value then we can group it by state.

Recommendations: As we can see state “RJ” has the highest order value so which is 163.94.

So if we need to increase the order value in every state like this then we need to take this as a case study

and we can analyze the factors that are making this happen and implement this strategy in other states

as well.

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

Here goes the query:

SELECT

ROUND(SUM(oi.freight_value), 2) AS `total_freight_value`,

ROUND(AVG(oi.freight_value), 2) AS `avg_freight_value`,

c.customer_state AS `state`

FROM

`Target.order_items` oi

JOIN

`Target.orders` o ON oi.order_id = o.order_id

JOIN

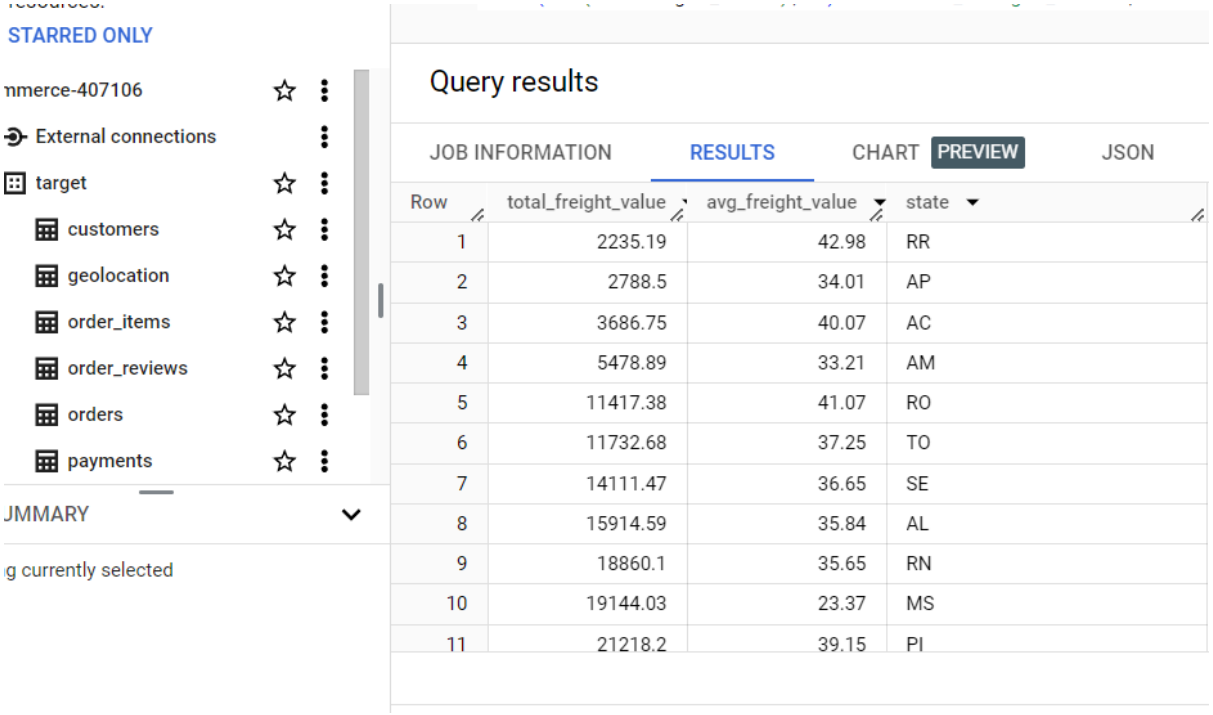
`Target.customers` c ON o.customer_id = c.customer_id

GROUP BY

`state`

ORDER BY 1;

Here is the screenshot:



Row	total_freight_value	avg_freight_value	state
1	2235.19	42.98	RR
2	2788.5	34.01	AP
3	3686.75	40.07	AC
4	5478.89	33.21	AM
5	11417.38	41.07	RO
6	11732.68	37.25	TO
7	14111.47	36.65	SE
8	15914.59	35.84	AL
9	18860.1	35.65	RN
10	19144.03	23.37	MS
11	21218.2	39.15	PI

Insights: we can clearly see that in the state SP avg_freight_value is the lowest so we can learn from here and what best works for this and then need to implement the same in other states as well.

Recommendations: NA

Q5(i)Find the no. of days taken to deliver each order from the order's purchase date as delivery time.

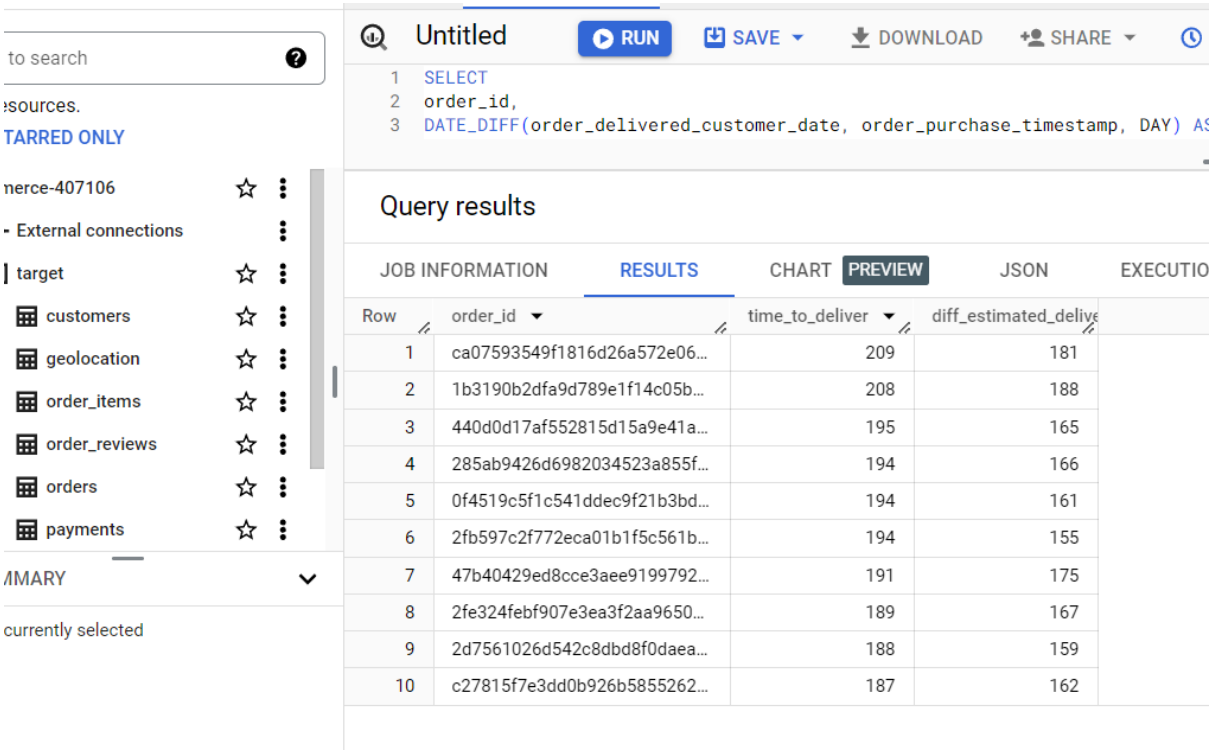
Also, calculate the difference (in days) between the estimated & actual delivery date of

an order.

Here is the query:

```
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_status = 'delivered'
order by 2 desc, 3 desc
```

Screenshot of the result:



Insights: From this data we are easily able to figure out the top orders where delivery happened really fast or late.

Recommendations: Again we can use this data to do further analysis like grouping by states and then figure out how many states are there which are delivering the order before estimated time and how many are after that. So when we do this kind of analysis we will definitely be able to figure out some regional pattern that affects the delivery timings. And then we can make the strategy accordingly to fix it.

Q5.Analysis based on sales, freight and delivery time.

(ii)

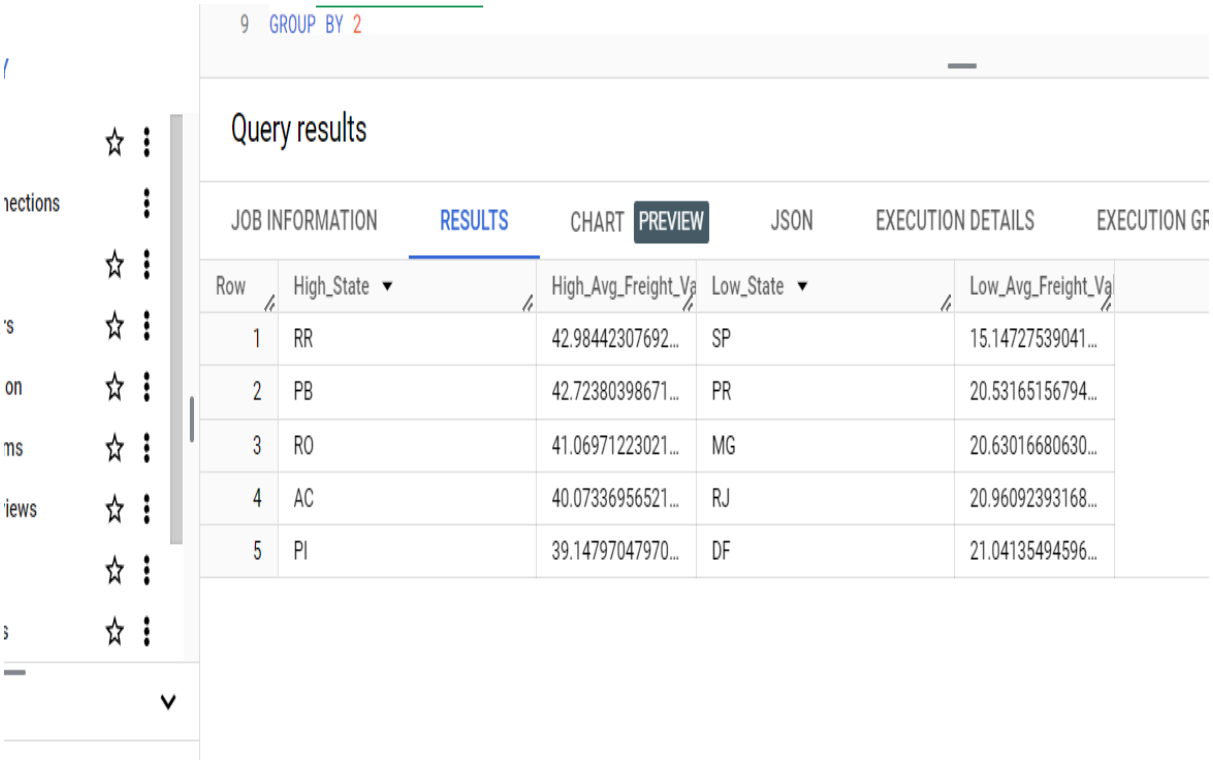
Here is the query:

```
-- Top 5 states with the highest average freight value
WITH TopStatesHigh AS (
SELECT AVG(oi.freight_value) as avg_freight_value,
c.customer_state as state,
ROW_NUMBER() OVER (ORDER BY AVG(oi.freight_value) DESC) as rank_high
FROM `Target.order_items` oi
JOIN `Target.orders` o ON oi.order_id = o.order_id
JOIN `Target.customers` c ON o.customer_id = c.customer_id
GROUP BY 2
LIMIT 5
)
-- Top 5 states with the lowest average freight value
, TopStatesLow AS (
SELECT AVG(oi.freight_value) as avg_freight_value,
c.customer_state as state,
ROW_NUMBER() OVER (ORDER BY AVG(oi.freight_value) ASC) as rank_low
FROM `Target.order_items` oi
JOIN `Target.orders` o ON oi.order_id = o.order_id
JOIN `Target.customers` c ON o.customer_id = c.customer_id
```

```
GROUP BY 2
LIMIT 5
)

-- Combine the results
SELECT
High.state AS High_State,
High.avg_freight_value AS High_Avg_Freight_Value,
Low.state AS Low_State,
Low.avg_freight_value AS Low_Avg_Freight_Value
FROM TopStatesHigh High
JOIN TopStatesLow Low ON High.rank_high = Low.rank_low;
```

Here is the screenshot of the result:



Row	High_State	High_Avg_Freight_Va	Low_State	Low_Avg_Freight_Va
1	RR	42.98442307692...	SP	15.14727539041...
2	PB	42.72380398671...	PR	20.53165156794...
3	RO	41.06971223021...	MG	20.63016680630...
4	AC	40.07336956521...	RJ	20.96092393168...
5	PI	39.14797047970...	DF	21.04135494596...

Insights: As per the query results we can see the top 5 lowest_avg

freight_value states and top 5 highest_avg freight_value states. RR is the highest with 42.98 and SP is the lowest with just 15.14

Recommendations: With this comparison we can take the positives and see the factors that are impacting this number. Then we can use the same model in all the states to save on cost.

Q5(iii).Find out the top 5 states with the highest & lowest average delivery time.

Ans:

Here is the query:

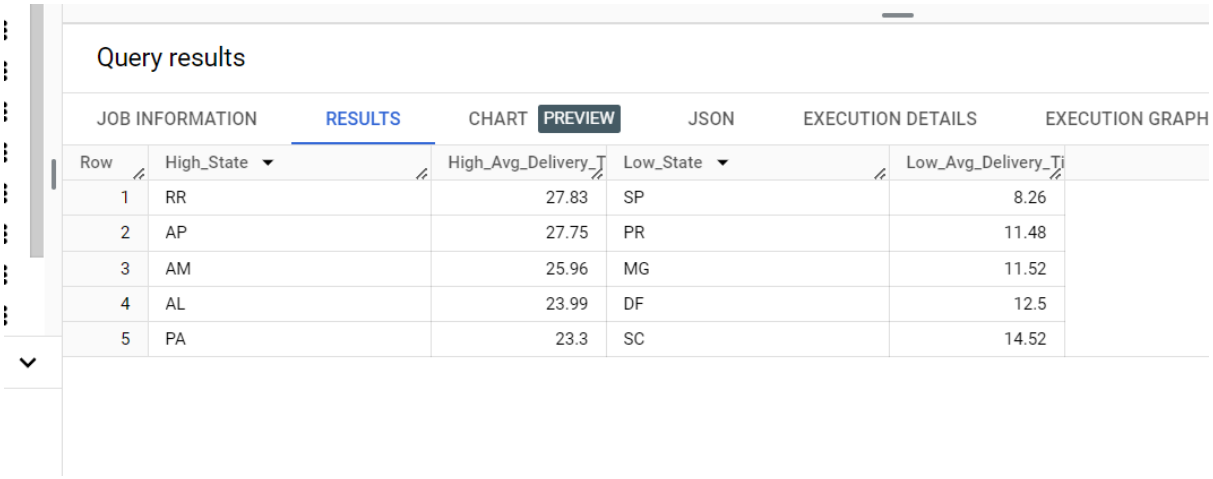
```
-- Top 5 states with the highest average delivery time
WITH TopStatesHigh AS (
SELECT
ROUND(AVG(DATE_DIFF(order_delivered_customer_date, order_purchase_timestamp, DAY)),
2) AS avg_delivery_time,
c.customer_state AS state,
ROW_NUMBER() OVER (ORDER BY AVG(order_delivered_customer_date -
order_purchase_timestamp) DESC) AS rank_high
FROM
`Target.order_items` oi
JOIN `Target.orders` o ON oi.order_id = o.order_id
JOIN `Target.customers` c ON o.customer_id = c.customer_id
GROUP BY 2
LIMIT 5
),
-- Top 5 states with the lowest average delivery time
TopStatesLow AS (
SELECT
ROUND(AVG(DATE_DIFF(order_delivered_customer_date, order_purchase_timestamp, DAY)),
2) AS avg_delivery_time,
c.customer_state AS state,
```

```
ROW_NUMBER() OVER (ORDER BY AVG(order_delivered_customer_date -
order_purchase_timestamp) ASC) AS rank_low
FROM
`Target.order_items` oi
JOIN `Target.orders` o ON oi.order_id = o.order_id
JOIN `Target.customers` c ON o.customer_id = c.customer_id
GROUP BY 2

LIMIT 5
)

SELECT
High.state AS High_State,
High.avg_delivery_time AS High_Avg_Delivery_Time,
Low.state AS Low_State,
Low.avg_delivery_time AS Low_Avg_Delivery_Time
FROM
TopStatesHigh High
JOIN TopStatesLow Low ON High.rank_high = Low.rank_low;
```

Screenshot of the result:



Query results

JOB INFORMATION | RESULTS | CHART | PREVIEW | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	High_State	High_Avg_Delivery_T	Low_State	Low_Avg_Delivery_Ti
1	RR	27.83	SP	8.26
2	AP	27.75	PR	11.48
3	AM	25.96	MG	11.52
4	AL	23.99	DF	12.5
5	PA	23.3	SC	14.52

Insights: As per the data it's very clear that “SP” is taking the

lowest delivery time but RR taking the highest time of 27 days.

Recommendations: Seems like it's really important to work on the logistics in the states where delivery is taking longer than expected. Otherwise we can lose the customers:

Q5(iv). Find out the top 5 states where the order delivery is really fast as compared to the estimated date of delivery.

Here is the query:

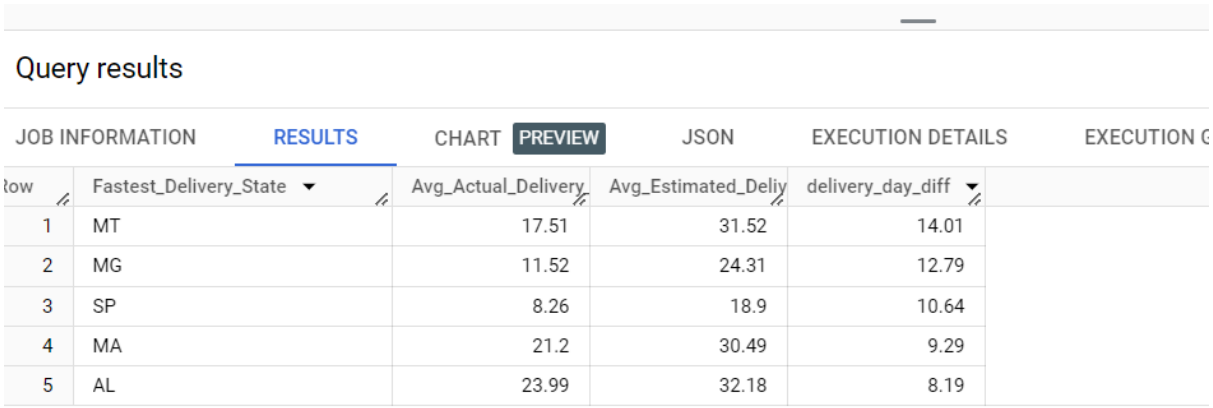
```
-- Top 5 states with the fastest delivery compared to estimated date
WITH FastestDeliveryStates AS (
SELECT
c.customer_state AS state,
ROUND(AVG(DATE_DIFF(order_delivered_customer_date, order_purchase_timestamp, DAY)),
2) AS avg_actual_delivery_time,
ROUND(AVG(DATE_DIFF(order_estimated_delivery_date, order_purchase_timestamp, DAY)),
2) AS avg_estimated_delivery_time,
ROW_NUMBER() OVER (ORDER BY AVG(order_delivered_customer_date -
order_purchase_timestamp) - AVG(order_estimated_delivery_date -
order_purchase_timestamp) DESC) AS rank_fast
FROM
`Target.order_items` oi
JOIN `Target.orders` o ON oi.order_id = o.order_id
JOIN `Target.customers` c ON o.customer_id = c.customer_id
GROUP BY 1
LIMIT 5
)

SELECT
state AS Fastest_Delivery_State,
```

```
avg_actual_delivery_time AS Avg_Actual_Delivery_Time,
avg_estimated_delivery_time AS Avg_Estimated_Delivery_Time,
round((avg_estimated_delivery_time - avg_actual_delivery_time),2)
as delivery_day_diff

FROM
FastestDeliveryStates
order by delivery_day_diff desc
```

Screenshot of the result:



Row	Fastest_Delivery_State	Avg_Actual_Delivery	Avg_Estimated_Deliy	delivery_day_diff
1	MT	17.51	31.52	14.01
2	MG	11.52	24.31	12.79
3	SP	8.26	18.9	10.64
4	MA	21.2	30.49	9.29
5	AL	23.99	32.18	8.19

Insights: As per the data “MT”is the state where delivery happens really fast. On an avg it's fast by 14 days from estimated timings. After that, “MG” also delivered 12 days faster than the estimate.

Recommendations: This data is really important to have a case study and figure out the reasons behind this much of fast deliveries. Once we get some pattern then we can do the same where delivery is being delayed.

Q6.Analysis based on the payments:

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

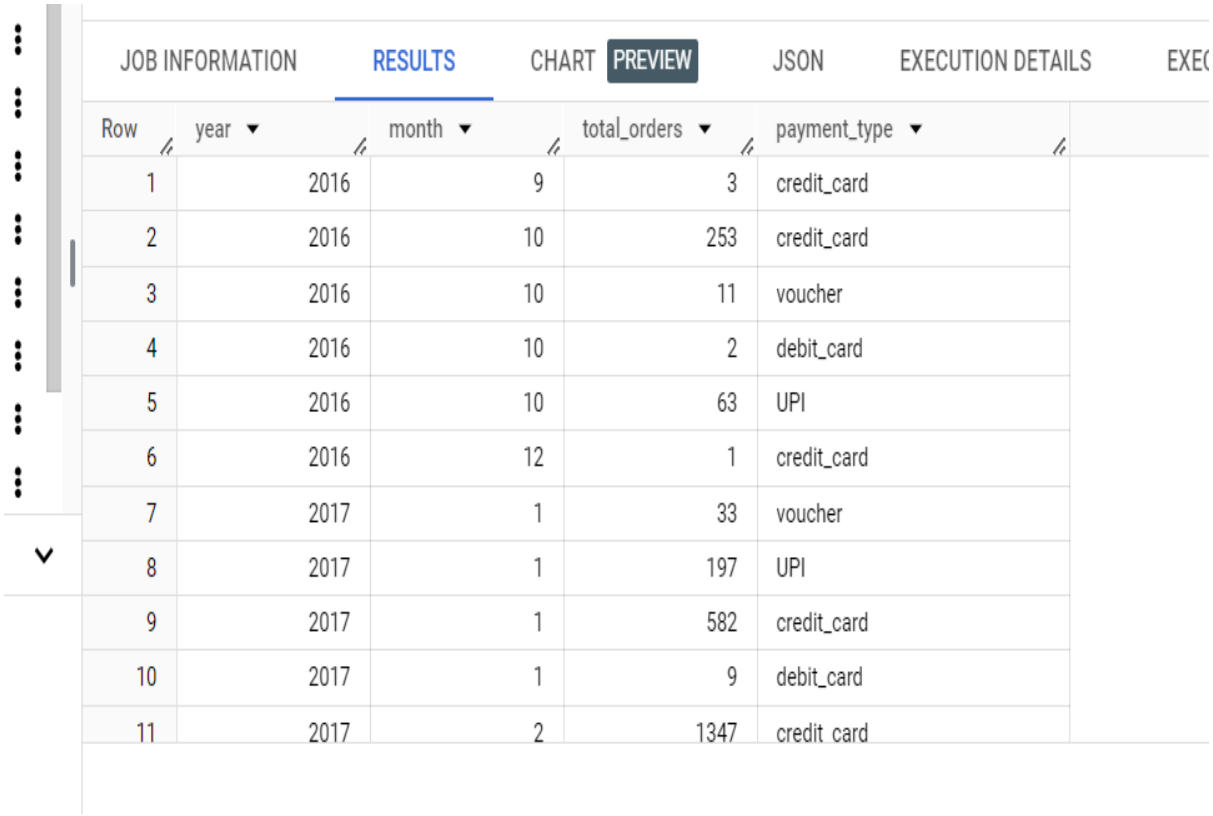
Ans:

Here is the query:

```
select
extract(year from o.order_purchase_timestamp) `year`,
extract(month from o.order_purchase_timestamp) `month`,
count(distinct o.order_id) `total_orders`,
p.payment_type `payment_type`

from `Target.orders` o
join `Target.payments` p
on o.order_id = p.order_id
group by `month`,`year`,`payment_type`
order by 1,2
```

Here is the screenshot of result:



JOB INFORMATION | RESULTS | CHART | PREVIEW | JSON | EXECUTION DETAILS | EXEC

Row	year	month	total_orders	payment_type
1	2016	9	3	credit_card
2	2016	10	253	credit_card
3	2016	10	11	voucher
4	2016	10	2	debit_card
5	2016	10	63	UPI
6	2016	12	1	credit_card
7	2017	1	33	voucher
8	2017	1	197	UPI
9	2017	1	582	credit_card
10	2017	1	9	debit_card
11	2017	2	1347	credit_card

Insights: By analyzing the data we can clearly see that most of the payments are made by using the credit cards.

Recommendations: So based on the data if we need to increase the no. of orders from other payment methods then we really need to run some offers on different methods so we can catch more customers.

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

Here is the query:

```
SELECT COUNT(DISTINCT o.order_id) as total_order,
p.payment_installments AS paid_installments
FROM `Target.payments` p
JOIN `Target.orders` o ON p.order_id = o.order_id
WHERE p.payment_installments >= 1
group by p.payment_installments
```

Here is the screenshot of result :

JOB INFORMATION | RESULTS | CHART | PREVIEW | JSON

Row	total_order	paid_installments
1	49060	1
2	12389	2
3	10443	3
4	7088	4
5	5234	5
6	3916	6
7	1623	7
8	4253	8
9	644	9
10	5315	10
11	23	11

Insights: to find the number of orders where at least one installment is paid we can count the distinct orders and then we can select a column of paid_installment and then we can group by paid_installments.

This will give us the grouped count of total orders as per installment paid.

Recommendation: As we can see that most of the orders are with 1 installment paid only so we have to analyze that most of the orders are 1 month old. That's why we have the most numbers of 1 installment or is there any other reason for that? Based on that we have to take the steps to collect the installments otherwise it will affect the revenue

Analyzing Expansion Prospects in Brazil for a Major US Retailer: SQL Business Case Study

Tool: Google BigQuery

Presented by: Avanti Raut

Problem Statement: 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 particular 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.

Business Insights:

- Based on the provided outcome, it's evident that, with the exception of the 'customer zip code prefix' column in the customer table, all other columns have a data type of string.
- Considering the dataset at hand, information spans from September 4th, 2016, to October 17th, 2018. This implies that the dataset encompasses Brazilian consumer data across a span of 772 days, which offers a substantial basis for thorough analysis. Additionally, the timestamps provided are in UTC, requiring a subtraction of 3 hours to align with the Brazilian time zone.
- The query outcome reveals that the dataset encompasses information from 27 different states and incorporates data from 4119 distinct cities situated within Brazil.
- Evidently, there is a noticeable upward trend evident when transition ing from 2016 to 2017. However, the possibility of aggregating data on a yearly basis is hindered due to the limited data availability for only three months in 2016, while comprehensive monthly data is accessible for the entirety of 2017.

  This prompted a comparison of order quantities on a monthly basis. The dataset clearly highlights a significant surge in order counts during the initial three months of 2017. Notably, a remarkable increase in orders is observed between December and January in both years. Furthermore, an intriguing observation emerges: a negative growth rate in order quantities occurs during the month of September for both the years 2017 and 2018.

- Comparing the data from the years 2016, 2017, and 2018 alongside each other, it becomes evident that there exists a monthly pattern within the data. As an example,

considering the case of March (row 3), a positive increase in order counts is observed. Conversely, for April, a decline in the number of orders is noticeable in both 2017 and 2018. Moreover, variations in growth rates between the two years are apparent in various other months as well. A

particularly noteworthy observation is the substantial drop in the number of orders during September across all years. Remarkably, this drop is particularly steep in 2018.

- Analyzing the query outcome, it becomes apparent that the peak number of orders occurs during the morning hours, specifically between 7 AM and 12 PM Brazilian time. Additionally, a substantial volume of orders is observed in the afternoon period (1 PM - 6 PM), which is reasonably expected.

However, a contrasting trend emerges during the night and dawn hours, with significantly fewer orders being placed. This suggests that consumers exhibit reluctance towards making purchases during these times.

- The results above illustrate the monthly count of orders placed for every state and year. This information could be valuable for delving into the specific order quantities per state in each month.
- The query outcome demonstrates that Sao Paulo contributes nearly 42 percent of our organization's customer base. This is unsurprising considering Sao Paulo's status as the most densely populated state in Brazil.

Aside from Sao Paulo, Rio de Janeiro and Minas Gerais are also noteworthy for generating substantial order volumes. Moreover, Rio Grande do Sul and Paraná's are among the top five states, each contributing around 5 percent of the total orders.

- The query outcome indicates that when contrasting January to August, there has been a surge of around 137 percent in the expenses associated with orders. This indicates an upward trend in people's expenditures on Target's items compared to the previous year. This phenomenon could be attributed to either an expansion in the customer base or an increase in spending per individual customer.
- The data extracted from the query reveals that Paraibas, Acre, and Alagoas are the states exhibiting higher average prices.

Conversely, Sao Paulo boasts the lowest average price across all states. An essential influencer for the states like Paran´a,Rio Grande do Sul with lower average prices is their predominant presence in the densely populated South and Southeast regions. Notably, Acre stands out due to its location in the tropical Amazon region, implying potentially elevated transportation costs for these states. As discernible from the subsequent query outcome, states with the highest average prices also tend to possess elevated average freight values, contributing to the overall price increase.

- In this context, it becomes evident that northern states such as Roraima, Paraibas, Rondônia, and Acre exhibit higher average freight costs, while states like Sao Paulo, Paraná's, and Rio de Janeiro have comparatively lower average freight expenses. This factor significantly influences the overall average prices of these states.
- The data retrieved from the aforementioned query provides a clear view of the delivery duration for individual products and the variance between their anticipated

and actual delivery times. The outcomes reveal a spectrum of deliveries occurring notably early, as well as instances where orders experience substantial delays. These inconsistencies could likely be attributed to the geographical context of certain northern states, situated in the Amazon tropical region or hilly terrains, which contrasts with the smoother delivery process in the southern states.

- Evidently, the foremost five states with the most elevated and least average freight values stand out prominently. As previously noted, northern states such as Roraima, Paraibas, Rondônia, and Acre demonstrate elevated average freight expenses, while states like Sao Paulo, Paraná's, and Rio de Janeiro showcase lowest average freight costs. This dynamic substantially shapes the overall average pricing trends within these states.
- The relationship between the five states with the shortest average delivery times and the lowest average freight costs is evident. This correlation is particularly notable as four out of these five states overlap. The states located in the southern region, namely Sao Paulo, Parana's, and Minas Gerais, exhibit notably swift average delivery times. In contrast, northern states like Roraima, Ampara, and Ama zonas experience extended average delivery times due to the challenging geographical conditions they face.
- It's puzzling that in the geographically challenging northern Brazilian states (Acre, Rondônia, Amazonas, Ampara), despite longer average delivery times, actual delivery is remarkably swift compared to estimates. This discrepancy can be explained by strategically setting conservative estimated delivery times to account for complex logistics, like harsh terrain or remote locations. This approach builds in a buffer for possible delays. When deliveries outpace estimates, it enhances customer satisfaction. This strategy aligns with prudent risk management and customer-centric objectives.

- The analysis depicts the monthly variation in the number of orders made through various payment options. The data makes it clear that credit card is the most preferred choice every month across all years. Following closely is UPI, which stands out as a noteworthy payment method due to its transaction convenience. In contrast, debit card usage as a payment method appears infrequent.
- The outcome reveals that the majority of orders are placed with a single payment installment. This preference could be attributed to the potential increase in interest payments for customers if they opt for multiple installments, motivating them to choose a single payment option.

Recommendations:

- The data reveals a consistent trend where demand experiences a surge during the early months of each year, specifically from January to March. How ever, this demand tapers off as the year progresses. Consequently, it might be advisable for the company to maintain higher inventory levels during these initial months. Con versely, a notable decline in demand is observed in September and December, indicating the need for reduced inventory during these months.

- Consequently, based on these findings, it is imperative to direct significant attention towards Sao Paulo, Rio de Janeiro, and Minas Gerais. Enhancements to customer amenities, post-sales services, and other provisions should be prioritized in these states. Additionally, the organization can strategize a marketing campaign tailored to the preferences and requirements of the inhabitants in these three states. In situations of parity, Sao Paulo should be given the highest priority.
- Consequently, the Brazilian market for Target is displaying growth compared to the previous year. Given the assumption of consistent spending per individual and an expanding customer base, Target could consider extending its operations within this region.
- Therefore, Target should prioritize its attention on states characterized by a substantial consumer base and lower average prices. This strategy would encourage greater purchasing due to the affordability of products. Moreover, considering that freight costs are lower in southern states like Sao Paulo, Paran´a,Rio Grande do Sul,Rio de Janeiro there is a favorable environment for efficient penetration and expansion.
- Given the regular utilization of credit cards, a strategy to encourage increased purchases could involve implementing a discount scheme for trans actions reaching a specific threshold or higher. Similarly, a comparable approach can be adopted for UPI transactions by collaborating with prominent UPI platforms to offer incentives for certain spending thresholds.
- There appears to be a correlation between higher interest rates and payment plans with more than one installment. To enhance customer attraction, particularly since credit card usage is predominant, strategies like introducing friction less credit options such as no-cost EMI or similar schemes could be implemented to streamline the credit process. 1