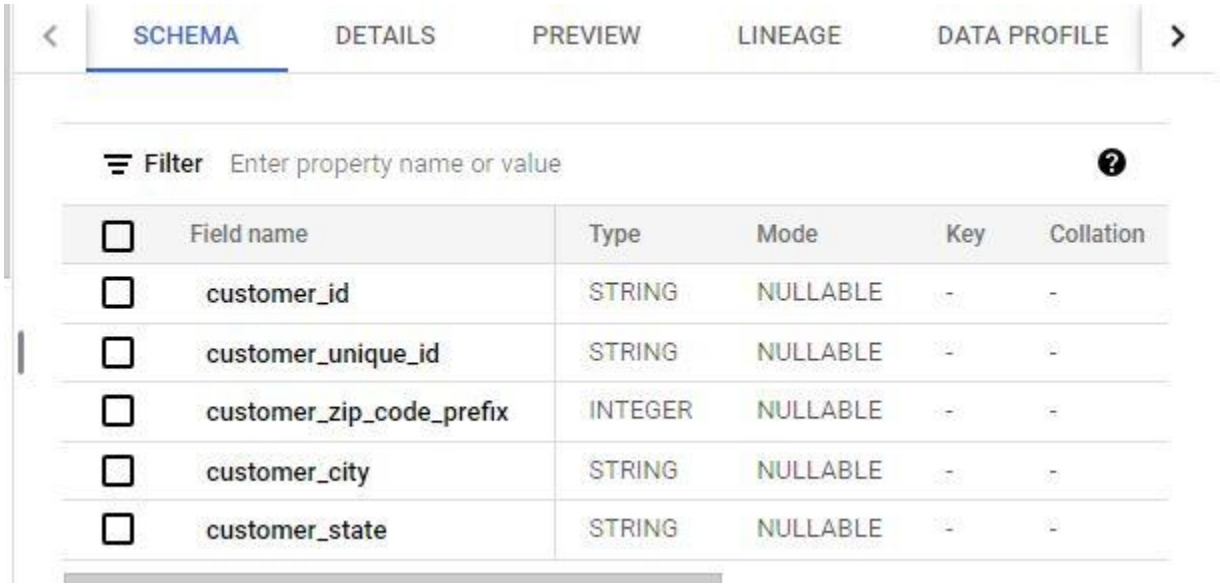
Business Case Study

Using BigQuery

Q1.1 Data types of all coloumns in the “customers” table.

	Type	Mode
customer_id	STRING	
customer_unique_id	STRING	
customer_zip_code_prefix	INTEGER	
customer_city	STRING	
customer_state	STRING	

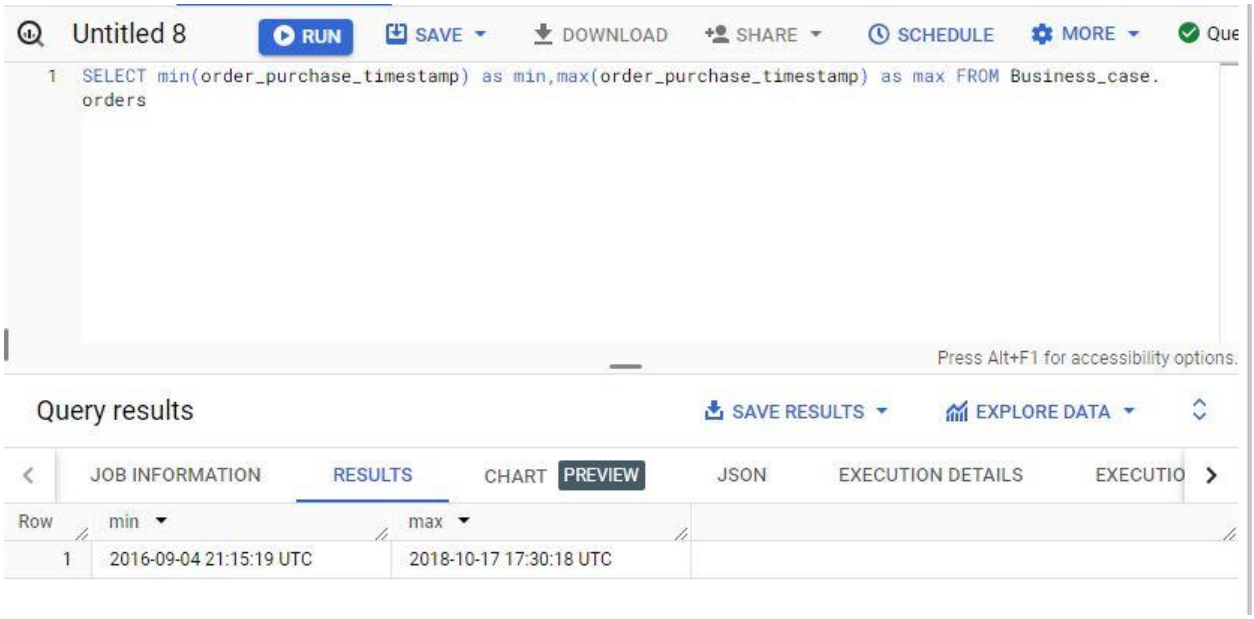


Recommendations : NA
Insights : NA

Q1.2. Time range between the orders were placed.

Query :

```
SELECT min(order_purchase_timestamp) as min,max(order_purchase_timestamp) as max FROM
 Business_case.orders
```



INSIGHTS :

The Orders are placed between “2016-09-04 21:15:19” and “2018-10-17 17:30:18”.

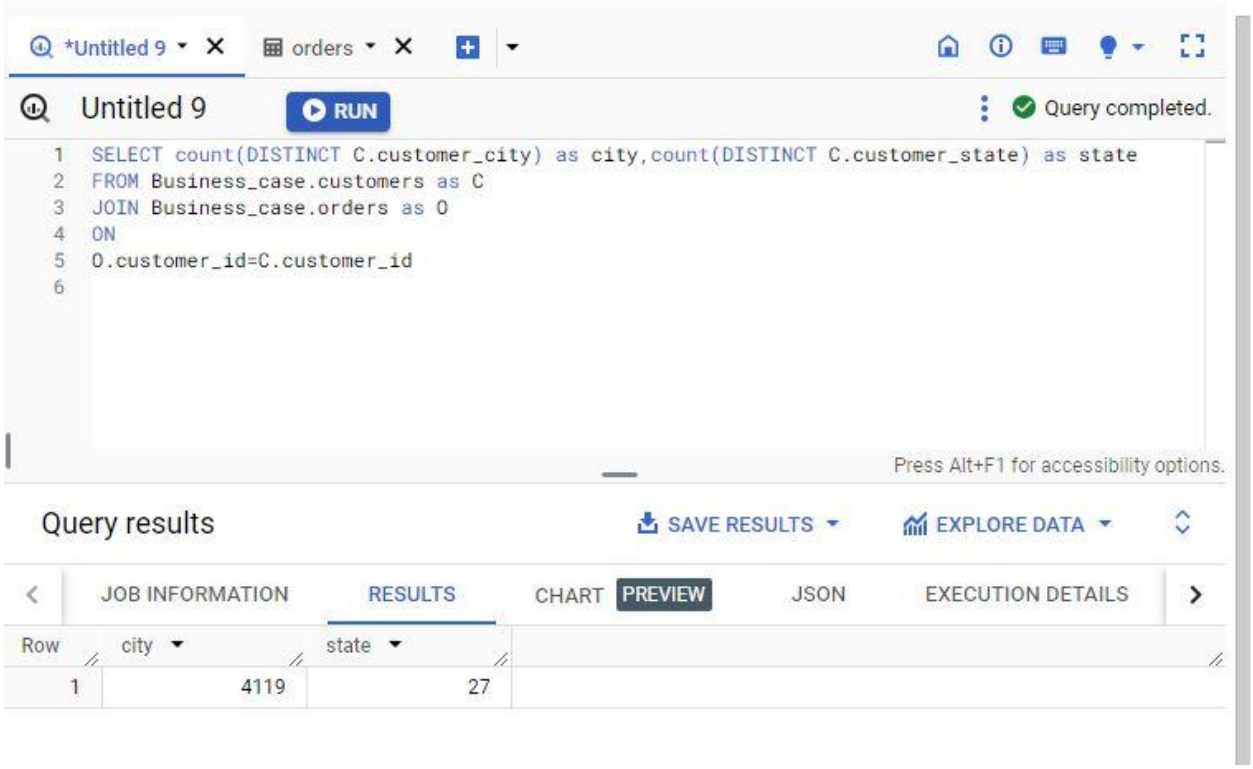
RECOMMENDATIONS :

1. Evaluate the trend, seasonal patterns using order data.
2. Determine peak hours and days in which orders are frequently placed.
3. Conduct marketing campaigns like discount offers, promotions etc during peak time and seasons.

Q1.3 : Count of Cities and States of Customers who ordered during the time period.

Query :

```
SELECT count(DISTINCT C.customer_city) as city,count(DISTINCT C.customer_state) as state
FROM Business_case.customers as C
JOIN Business_case.orders as O  ON O.customer_id=C.customer_id
```



INSIGHTS :

Customers from 4119 Cities across 27 States have placed the order during this time range.

RECOMMENDATIONS :

1.Analyze the distribution of orders across different states. Identify regions with the highest and lowest order volumes.
2. Improve inventory and shipping in regions with highest volume which would enhance customer experience.
3. Try to expand business in cities with lowest orders.
4.Understand product preferences and buying behavior in different states and cities. Consider adapting product offerings or marketing messages based on regional preferences and trends.
5. Take into account cultural differences and preferences in marketing and product positioning. Do branding according to the cultural diversity in different states and regions.

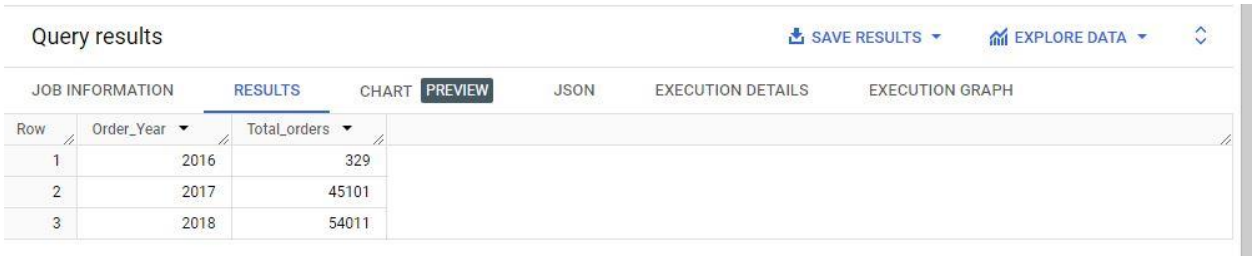
Q2.1 :

Growing trend in the number of orders placed in the past years.

Query :

```
SELECT EXTRACT (YEAR FROM order_purchase_timestamp) as Order_year,
count(order_id) as Total_orders FROM Business_case.orders

GROUP BY Order_year
ORDER BY Order_year
```



Row	Order_Year	Total_orders
1	2016	329
2	2017	45101
3	2018	54011

INSIGHTS :

We observe a consistent upward trend in the number of orders each year. Notably, the company has experienced a remarkable 13,708.51% increase in the number of orders during its initial year, indicating a strong establishment in the market. This substantial growth from 2017 to 2018 underscores the company's significant progress in establishing itself and expanding its market presence.

RECOMMENDATIONS:

Given the upward trajectory in order numbers, it is advisable for the company to enhance its operational capacity. This includes augmenting inventory levels, optimizing logistics processes, and expanding human resources. Key areas for recruitment should focus on increasing the staff in critical areas such as customer service, transportation, and warehouse operations. By strategically addressing these aspects, the company can better accommodate the growing demand and ensure a seamless experience for both customers and internal operations.

Q2.2 :

Monthly seasonality in terms of the no. of orders being placed :

Query :

```
SELECT EXTRACT (Month FROM order_purchase_timestamp) as Order_month,
count(order_id) as Total_orders FROM Business_case.orders

GROUP BY Order_month
ORDER BY Order_month
```

Query results

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

Row	Order_month	Total_orders
1	1	8069
2	2	8508
3	3	9893
4	4	9343
5	5	10573
6	6	9412
7	7	10318
8	8	10843
9	9	4305
10	10	4959
11	11	7544
12	12	5674

INSIGHTS :

The analysis reveals a notable surge in the number of orders, particularly during the months of May, July, and August. This pattern suggests a seasonal trend in order placement. One explanation for this trend is that these months coincide with the winter season in Brazil. During this period, there is a trend for increased reliance on e-commerce, possibly attributed to the challenges of purchasing items from local stores amidst the winter conditions.
We can also see that the sales are at its lowest during the month of September to December. This clearly indicates that there is seasonality in number of orders placed by the customers.

RECOMMENDATIONS :

To boost sales during the period of September to December, it is recommended to implement a strategic discount program aimed at attracting more customers. Offering discounts during this traditionally slower season has the potential to stimulate customer interest, increase sales, and consequently, enhance overall revenue for the company. By strategically promoting and communicating these discounts, the company can effectively capitalize on the lower sales period, encouraging customer engagement and fostering a positive impact on its financial performance during this season.

In anticipation of the peak season from May to August, proactive measures should be taken to ensure optimal preparedness. This includes augmenting inventory levels, expanding the number of available stocks, and disseminating timely alerts to company employees about the upcoming peak period. By enhancing inventory and stock levels, the company can meet heightened demand efficiently, while ensuring that employees are well-informed and prepared contributes to seamless operations during this crucial time. Such preparatory steps are essential for capitalizing on the increased demand and maintaining a high level of service excellence throughout the peak months.

Q2.3 :

During what time of the day, do the Brazilian customers mostly place their orders?

Query:

```
SELECT count(order_id) as Total_orders,

CASE

WHEN EXTRACT(TIME FROM order_purchase_timestamp) BETWEEN "00:00:00" AND "06:00:00" THEN "Dawn"

WHEN EXTRACT(TIME FROM order_purchase_timestamp) BETWEEN "07:00:00" AND "12:00:00" THEN
"Mornings"

WHEN EXTRACT(TIME FROM order_purchase_timestamp) BETWEEN "13:00:00" AND "18:00:00" THEN
"Afternoon"

WHEN EXTRACT(TIME FROM order_purchase_timestamp) BETWEEN "19:00:00" AND "23:00:00" THEN
"Night"

ELSE "Z"

END AS Order_time

FROM Business_case.orders

GROUP BY Order_time

ORDER BY Order_time
```

Query results

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

Row	Total_orders	Order_time
1	32370	Afternoon
2	4740	Dawn
3	21738	Mornings
4	24209	Night
5	16384	Z

INSIGHTS :

From the data we can understand that most orders are placed in the “Afternoon” followed by “Night” .
The least number of orders are placed at “Dawn”.
There is significant amount of orders placed during “Morning” time as well.
This indicates that customers prefer to place orders between 1PM to 11PM which is the “peak time” in a day.

RECOMMENDATIONS:

1. Optimize website features and speed during the peak time to manage the high volume.
2. Customer service staffing can be optimized between 1PM and 11PM and can be reduced at Dawn.
3. Ordering promoting strategies like discounts, promotional offers, Cash back offers, fastest shipping etc can be provided to increase the order volume during Dawn time, thereby increasing business.
4. Marketing strategies should be done during the peak time, maximize advertisement and promotional activities for maximum visibility and engagement.
5. Adjust inventory levels based on order patterns throughout the day. Ensure that popular products are well-stocked during peak hours to meet increased demand and minimize the risk of stockouts
6. Implement personalized recommendations and targeted marketing messages based on the time of day. For example, suggest dinner-related items in the late afternoon and evening, and breakfast items in the morning.

Q3.1 : **Get the month on month no. of orders placed in each state.**

Query :

```
SELECT
count(order_id) as Total_orders,customers.customer_state as customer_state,

EXTRACT(Month FROM order_purchase_timestamp) as Month
FROM
Business_case.orders
JOIN
Business_case.customers
ON
orders.customer_id=customers.customer_id
GROUP BY customer_state,Month
ORDER BY customer_state,Month
```

Query results

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

Row	Total_orders	customer_state	Month
1	8	AC	1
2	6	AC	2
3	4	AC	3
4	9	AC	4
5	10	AC	5
6	7	AC	6
7	9	AC	7
8	7	AC	8
9	5	AC	9
10	6	AC	10
11	5	AC	11
12	5	AC	12

INSIGHTS :

The state of "MG" records the highest number of orders, with March having the highest order count among all states, totaling 1237 orders. Over the course of 12 months, a total of 11635 orders are placed in the state "MG."
In contrast, the state of "AP" has the lowest number of orders, with only 68 orders placed throughout the entire year. Notably, September is the month with the lowest order count for the state of "AP," having only 2 orders placed during that month, marking the lowest among all states.

RECOMMENDATIONS :

1. Increase Marketing strategies like Advertisement in the state of “AP” for maximum visibility.
2. Provide promotional offers, discounts and cashbacks in the state of “AP” during local festival seasons which would attract more customers in this region.
3. Optimize inventory and transportation in the State of “MG” for better customer experience.

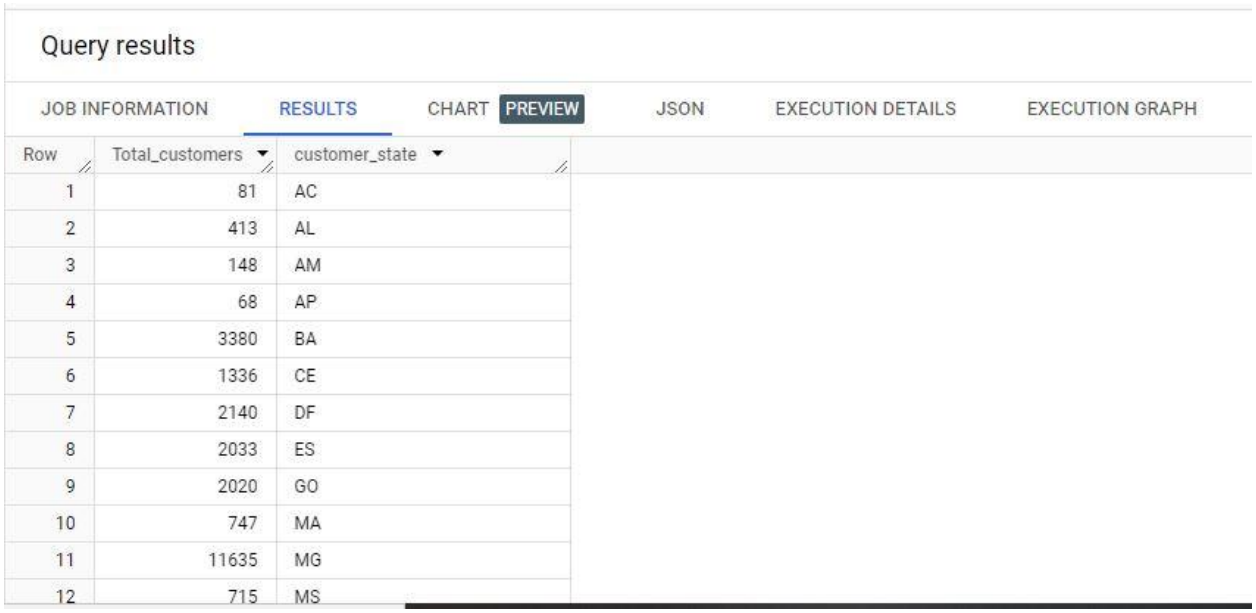
Q3.2: How are the customers distributed across all the states?

Query :

```
SELECT

count(customer_id) as Total_customers,customer_state

FROM
Business_case.customers
GROUP BY customer_state
ORDER BY customer_state
```



Query results

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

Row	Total_customers	customer_state
1	81	AC
2	413	AL
3	148	AM
4	68	AP
5	3380	BA
6	1336	CE
7	2140	DF
8	2033	ES
9	2020	GO
10	747	MA
11	11635	MG
12	715	MS

INSIGHTS :

The state of "SP" boasts the highest number of customers, totaling 41,746, while the state of "RR" has the fewest customers, with only 46.
Among the states, "MG," "RJ," and "SP" are the only ones with customer counts exceeding 10,000. On the other hand, "AP," "AC," and "RR" have the lowest number of customers, each having less than 100 customers in their respective states.

RECOMMENDATIONS:

1. Expand the business in the States of “AP”,”AC” and “RR” by doing good marketing strategies and providing promotional offers during local festival seasons.

2. Improve shipping, inventory in the States of “MG”,”RJ” and “SP” to give better customer experience.

3. Expand business in other States as well using points mentioned in the first recommendation.

Q4.1: Get the % increase in the cost of orders from year 2017 to 2018 (include months between Jan to Aug only).

Query :

```
    WITH OrderCosts AS (
  SELECT
    EXTRACT(YEAR FROM o.order_purchase_timestamp) AS order_year,
    EXTRACT(MONTH FROM o.order_purchase_timestamp) AS order_month,
    p.payment_value
  FROM
    Business_case.orders o
    JOIN Business_case.payments p ON o.order_id = p.order_id
  WHERE
    EXTRACT(YEAR FROM o.order_purchase_timestamp) IN (2017, 2018)
    AND EXTRACT(MONTH FROM o.order_purchase_timestamp) BETWEEN 1 AND 8
)
SELECT
  order_year,
  order_month,
  SUM(payment_value) AS total_cost,
  (SUM(payment_value) - LAG(SUM(payment_value), 1) OVER (ORDER BY order_year, order_month)) /
LAG(SUM(payment_value), 1) OVER (ORDER BY order_year, order_month) * 100 AS
percentage_increase
FROM
  OrderCosts
GROUP BY
  order_year, order_month
ORDER BY
  order_year,order_month
```

Query results

SAVE RESULTS | EXPLORE DATA

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

Row	order_year	order_month	total_cost	percentage_increase
1	2017	1	138488.0399999...	null
2	2017	2	291908.0099999...	110.7821079712...
3	2017	3	449863.6000000...	54.11142708965...
4	2017	4	417788.0300000...	-7.13006564656...
5	2017	5	592918.8200000...	41.91857531198...
6	2017	6	511276.3800000...	-13.7695814749...
7	2017	7	592382.9200000...	15.86354135898...
8	2017	8	674396.3200000...	13.84465980214...
9	2018	1	1115004.180000...	65.33366908051...
10	2018	2	992463.3400000...	-10.9901686646...

INSIGHTS :

The percentage increase is 0 to negative in most of the months from February to August in the year 2018 which indicates that the cost of orders were drastically reduced during this period.

The highest percentage increase is in month of February 2017 which indicates that the cost of orders went more than double in this month.

RECOMMENDATIONS:

1. Analysis should be done on why the cost of order went 110% high in Feb 2017 and proper business decisions should be taken according.
2. The cost of orders has not increased in some months where percentage increase is 0 and has reduced in some months where it is negative. Analysis should be done on this as well as it is a scope for business.

Q4.2 : Total & Average value of order price for each state.

Query :

```
SELECT
  customers.customer_state,
  SUM(order_items.price) AS total_order_price,
  AVG(order_items.price) AS average_order_price

FROM
Business_case.customers
JOIN
Business_case.orders
ON
customers.customer_id = orders.customer_id
JOIN
Business_case.order_items
ON
orders.order_id = order_items.order_id
WHERE
customers.customer_state IS NOT NULL
GROUP BY
customer_state
ORDER BY
customer_state;
```

Query results

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

Row	customer_state	total_order_price	average_order_price
1	AC	15982.94999999...	173.7277173913...
2	AL	80314.81	180.8892117117...
3	AM	22356.84000000...	135.4959999999...
4	AP	13474.29999999...	164.3207317073...
5	BA	511349.9900000...	134.6012082126...
6	CE	227254.7099999...	153.7582611637...
7	DF	302603.9399999...	125.7705486284...
8	ES	275037.3099999...	121.9137012411...
9	GO	294591.9499999...	126.2717316759...
10	MA	119648.2199999...	145.2041504854...
11	MG	1585308.029999...	120.7485741488...
12	MS	116812.6399999...	142.6283760683...

INSIGHTS :

SP, RJ and MG are the top 3 States with highest Total order price indicating that most business happened here.

AC, AP and RR are the States with lowest Total order price.

PB, AL and AC are the States with highest Average order price

RS, PR and SP are the States with lowest Average order price.

RECOMMENDATIONS:

1. High average order price indicates that customers in this region tend to make high value orders. So visibility of high valued items in the website home page should be done in these regions.
2. High Total order price indicates the high volume of orders so inventory should be optimized so that products stockouts doesn't happen.

Q4.3 : Total & Average value of order freight for each state.

Query :

```
        SELECT
  customers.customer_state,
  SUM(order_items.freight_value) AS total_freight_price,
  AVG(order_items.freight_value) AS average_freight_price
FROM
  Business_case.customers
JOIN
  Business_case.orders
ON
  customers.customer_id = orders.customer_id
JOIN
  Business_case.order_items
ON
  orders.order_id = order_items.order_id
WHERE
  customers.customer_state IS NOT NULL
GROUP BY
  customer_state
ORDER BY
  customer_state;
```

Query results

JOB INFORMATION | RESULTS | CHART | PREVIEW | JSON | EX

Row	customer_state	total_freight_price	average_freight_price
1	AC	3686.749999999...	40.07336956521...
2	AL	15914.58999999...	35.84367117117...
3	AM	5478.889999999...	33.20539393939...
4	AP	2788.500000000...	34.00609756097...
5	BA	100156.6799999...	26.36395893656...
6	CE	48351.58999999...	32.71420162381...
7	DF	50625.49999999...	21.04135494596...
8	ES	49764.59999999...	22.05877659574...
9	GO	53114.97999999...	22.76681525932...
10	MA	31523.77000000...	38.25700242718...
11	MG	270853.4600000...	20.63016680630...

INSIGHTS :

Highest Total freight price is in SP,RJ and MG.

Lowest Total freight price is in AC, AP and RR.

Highest Average freight price is in RR, PB and RO. This implies that, on average, individual freight transactions in these states have higher costs compared to other states.

Lowest Average Freight price is in MG, PR and SP. This indicates that, on average, individual freight transactions in these states have lower costs.

RECOMMENDATIONS :

1. "SP," "RJ," and "MG" these states contribute the most to the overall sum of freight costs. It could be due to a high volume of shipments, longer distances. Hence shipping department should be optimized, increasing the number of carriers in the region can be done as the volume of shipping is high.
2. There is a possibility of expansion in the States of "AC," "AP," and "RR" as Freight price is low in this States which could be because of shorter shipping distances, lower shipping volumes, or more cost-effective transportation methods. When Freight price is low the total order cost will also be low.

Q5.1 :

Number 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 order.

Query :

```
SELECT
  order_id,
  DATE_DIFF(DATETIME(order_delivered_customer_date), DATETIME(order_purchase_timestamp), DAY)
as delivery_time,
  DATE_DIFF(DATETIME(order_delivered_customer_date), DATETIME(order_estimated_delivery_date),
DAY) as delivery_difference
FROM
  Business_case.orders;
```

Query results

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

Row	order_id	delivery_time	delivery_difference
1	1950d777989f6a877539f5379...	30	12
2	2c45c33d2f9cb8ff8b1c86cc28...	31	-29
3	65d1e226dfaeb8cdc42f66542...	36	-17
4	635c894d068ac37e6e03dc54e...	31	-2
5	3b97562c3aee8bdedcb5c2e45...	33	-1
6	68f47f50f04c4cb6774570cfde...	30	-2
7	276e9ec344d3bf029ff83a161c...	44	4
8	54e1a3c2b97fb0809da548a59...	41	4
9	fd04fa4105ee8045f6a0139ca5...	37	1
10	302bb8109d097a9fc6e9cefc5...	34	5
11	66057d37308e787052a32828...	39	6

INSIGHTS :

Delivery time is high, in most of the cases it is above 30 days but in some cases delivery difference is negative indicating that the packages are delivered after estimated delivery date.

RECOMMENDATIONS :

1. Efforts should be made to reduce the delivery time, high delivery time will reduce customer satisfaction. Better shipping facilities should be implemented.
2. Effective shipping carriers should be identified and tied up with the company.
3. Delivering packages before the estimated delivery date will increase customer satisfaction and customer experience.

Q5.2 : Top 5 states with the highest & lowest average freight value.

Query :

```
SELECT
  seller_state,
  AVG(freight_value) AS Average
FROM
  Business_case.order_items
JOIN
  Business_case.Sellers
ON
  order_items.seller_id = Sellers.seller_id
GROUP BY
  seller_state
ORDER BY
  Average ASC
LIMIT 5;
```

Query results

JOB INFORMATION | RESULTS | CHART | PREVIEW | JSON

Row	seller_state	Average
1	SP	18.45221266585...
2	PA	19.38874999999...
3	RJ	19.47486508924...
4	DF	20.57181312569...
5	PR	22.72096874639...

INSIGHTS :

The highest Average freight value among top 5 States is “PR” and lowest is “SP”

RECOMMENDATIONS :

1. Better freight carrier should be identified in “PR” which could reduce freight value which in turn will reduce total order cost, reduced order cost will attract more customers to the company.
2. Marketing can be done in “SP” based on low freight value

Q5.3 : Top 5 states with the highest & lowest average delivery time

Query :

```
SELECT

  customers.customer_state,
  ROUND (AVG(DATE_DIFF(DATETIME(order_delivered_customer_date),
DATETIME(order_purchase_timestamp), DAY))) AS Average

FROM
 Business_case.customers
JOIN
 Business_case.orders
ON
 customers.customer_id=orders.customer_id
GROUP BY
 customer_state
ORDER BY Average ASC
LIMIT 5
```

Query results

JOB INFORMATION | RESULTS | CHART | PREVIEW

Row	customer_state	Average
1	SP	9.0
2	MG	12.0
3	PR	12.0
4	DF	13.0
5	RS	15.0

INSIGHTS :

Highest Average delivery time is in the state of “RS” among top 5 states and lowest is in “SP”

RECOMMENDATIONS :

1. Shipping should be made effective in “RS” as high delivery time will affect customer experience.
2. Inventory should be optimized in “RS” so that there is no delay in shipping out the item, there by improving delivery time.

Q5.4 :

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

Query :

```
SELECT
  customer_state,
  Average_estimated_time,
  Average_delivery_time,
  (Average_estimated_time-Average_delivery_time) AS Difference
FROM
(SELECT
  customers.customer_state,
  ROUND (AVG(DATE_DIFF(DATETIME(order_delivered_customer_date),
DATETIME(order_purchase_timestamp), DAY))) AS Average_delivery_time,
  ROUND (AVG(DATE_DIFF(DATETIME(order_estimated_delivery_date),
DATETIME(order_purchase_timestamp), DAY))) AS Average_estimated_time,
  FROM
 Business_case.customers
JOIN
 Business_case.orders
ON
 customers.customer_id=orders.customer_id
GROUP BY
 customer_state) AS Subquery

ORDER BY Average_delivery_time,Average_estimated_time
LIMIT 5
```

Query results

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

Row	customer_state	Average_estimated_t	Average_delivery_tim	Difference
1	SP	20.0	9.0	11.0
2	MG	25.0	12.0	13.0
3	PR	25.0	12.0	13.0
4	DF	25.0	13.0	12.0
5	SC	26.0	15.0	11.0

INSIGHTS :

The difference between average delivery date and average estimated delivery date is highest in “MG” and “PR”. Lowest is in “SP” and “SC”. Which means “MG” and “PR” is having fastest delivery and “SP” and “SC” is having slowest delivery.

RECOMMENDATIONS :

1. Effective transportation and shipping should be done in “SP” and “SC” to speed up the delivery.
2. Inventory should be optimized so that there is no lag or delay in shipping the item from the warehouse.
3. Advertisements can be done in “MG” and “PR” based on fastest shipping speed in the region.

Q6.1 : Month on month no. of orders placed using different payment types.

Query :

```
SELECT
 payments.payment_type,
 count(orders.order_id) AS Total_orders,
EXTRACT(Month FROM order_purchase_timestamp) AS Month
FROM
 Business_case.orders
JOIN
 Business_case.payments
ON
 orders.order_id=payments.order_id
GROUP BY
 Month,payment_type
ORDER BY
 Month,Total_orders ASC
```

Query results

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

Row	payment_type	Total_orders	Month
1	debit_card	118	1
2	voucher	477	1
3	UPI	1715	1
4	credit_card	6103	1
5	debit_card	82	2
6	voucher	424	2
7	UPI	1723	2
8	credit_card	6609	2
9	debit_card	109	3
10	voucher	591	3
11	UPI	1942	3
12	credit_card	7707	3

INSIGHTS :

From the data acquired we can see that most of the customers prefer using credit cards for buying items in every month. Customers least prefer debit cards as a payment methods for purchasing items. There is a significant use of UPI transactions as well.

RECOMMENDATIONS :

1. Business deals can be made with banks or debit card companies to offer promotions or discounts while using debit cards as a payment method. Thereby promoting the use of debit cards.

Q6.2 : Number of orders placed on the basis of the payment installments that have been paid.

Query :

```
SELECT
 payment_installments,
 count(order_id) AS Total_orders
FROM
 Business_case.payments
GROUP BY
 payment_installments
```

Query results

JOB INFORMATION | RESULTS | CHART | PREVIEW

Row	payment_installment	Total_orders
1	0	2
2	1	52546
3	2	12413
4	3	10461
5	4	7098
6	5	5239
7	6	3920
8	7	1626
9	8	4268
10	9	644
11	10	5328
12	11	23

INSIGHTS :

We can see that most customers prefer lesser number of payment installments as Total orders are high when payment installments are less.

RECOMMENDATIONS :

1. Provide educational content or guides on the benefits of different payment installment options. Help customers understand how choosing fewer installments can simplify their financial planning and reduce the overall repayment period.
2. Emphasize any cost savings or discounts associated with choosing fewer payment installments.
3. Ensure that customers have easy access to information about any additional fees or interest rates associated with longer payment plans.