Business Case: Target SQL

- Kishore.S

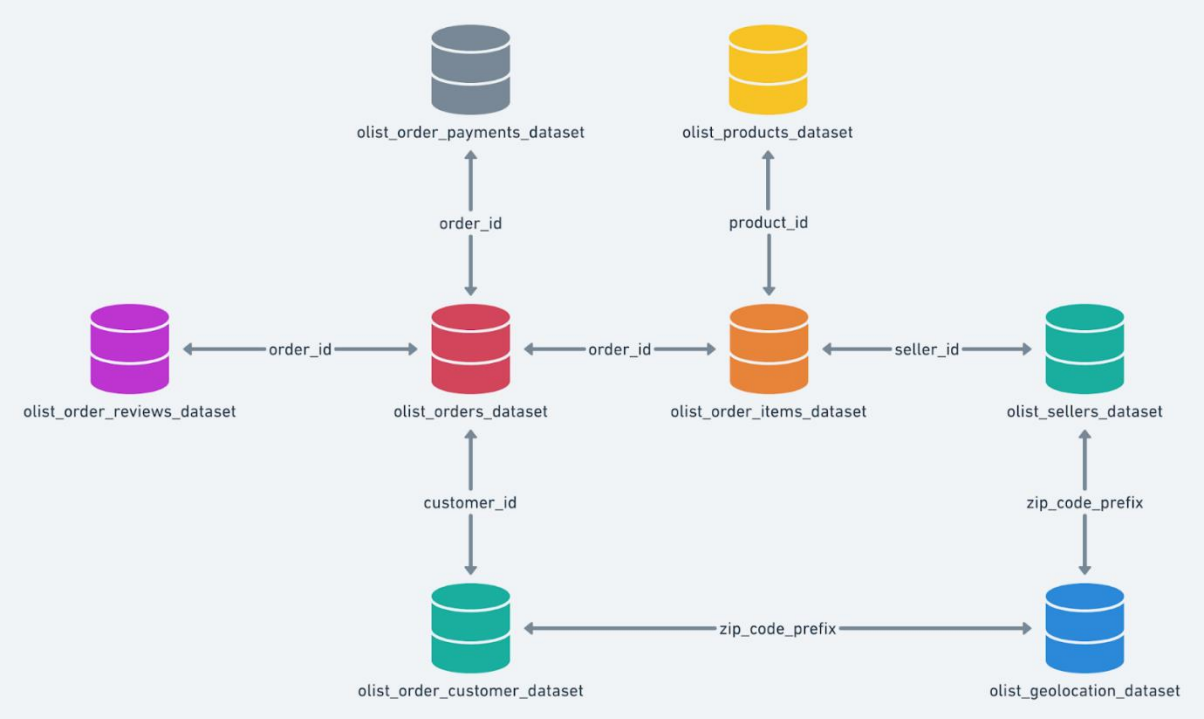
Context :

- 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 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.

Dataset :

- customers.csv
- sellers.csv
- order_items.csv
- geolocation.csv
- payments.csv
- reviews.csv
- orders.csv
- products.csv

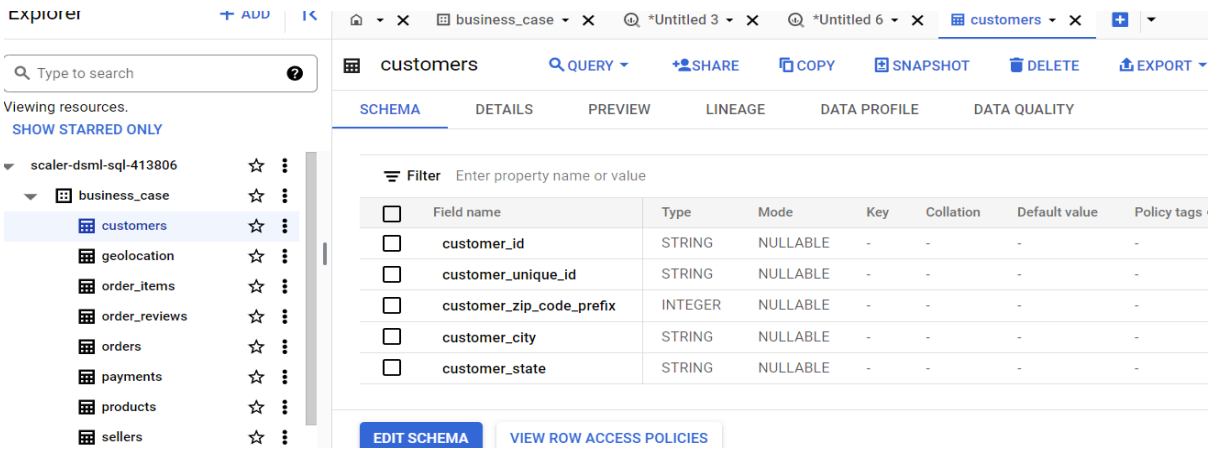
Schema :



1. 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

By clicking on each table,we can find out the data type of the columns



Alternative way:

```
SELECT
 column_name,
 data_type
FROM
 business_case.INFORMATION_SCHEMA. COLUMNS
ORDER BY column_name
```

Query results

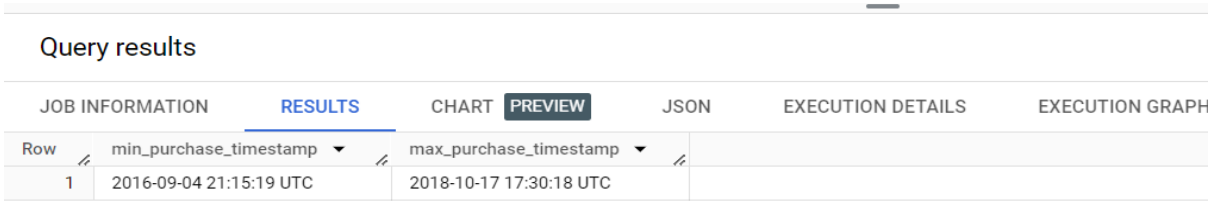
JOB INFORMATION | RESULTS | CHART | PREVIEW | JSON

Row	column_name	data_type
1	customer_city	STRING
2	customer_id	STRING
3	customer_id	STRING
4	customer_state	STRING
5	customer_unique_id	STRING
6	customer_zip_code_prefix	INT64
7	freight_value	FLOAT64

Overall, the data types seem appropriate for their respective fields. The STRING data types for customer-related information are suitable for textual representations, while INT64 is appropriate for numerical data like zip code prefixes.

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

```
SELECT
MIN(order_purchase_timestamp) AS
min_purchase_timestamp,
MAX(order_purchase_timestamp) AS
max_purchase_timestamp
FROM `business_case.orders`
```



The dataset spans from September 4, 2016 at 21:15:19 UTC to October 17, 2018, at 17:30:18 UTC. This 2-year and 1-month timeframe provides ample data for analysis, enabling insights into long-term trends,seasonality, and changes in customer behavior.

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

```
SELECT
 COUNT(DISTINCT customer_city) AS num_cities,
 COUNT(DISTINCT customer_state) AS num_states
FROM
 `business_case.customers`
```

Query results

JOB INFORMATION | RESULTS | CHART | PREVIEW | JS

Row	num_cities	num_states
1	4119	27

The dataset comprises 4,119 distinct cities spanning across 27 states, indicating extensive market coverage that could impact marketing and delivery approaches. Essentially, the dataset mirrors a varied customer demographic, encompassing numerous cities and states. Leveraging this data, we can customize strategies according to regional nuances, thus amplifying customer interaction.

2. Indepth Exploration :

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

```
SELECT
 EXTRACT(YEAR
 FROM
  order_purchase_timestamp) AS order_year,
 COUNT(*) AS num_orders
FROM
 `business_case.orders`
GROUP BY
 order_year
ORDER BY
 order_year;
```

Query results

JOB INFORMATION | RESULTS | CHART | PREVIEW

Row	order_year	num_orders
1	2016	329
2	2017	45101
3	2018	54011

During 2016, there were 329 orders, marking the outset of data collection. Subsequent years exhibited remarkable growth, with 2017 experiencing a notable surge to 45,101 orders, followed by 2018 sustaining the upward trajectory with 54,011 orders. This data can guide us in projecting future order volumes and allocating resources accordingly. Additionally, we can pinpoint peak seasons or periods of heightened demand for proactive readiness.

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

```
SELECT
 EXTRACT(MONTH
 FROM
  order_purchase_timestamp) AS order_month,
 COUNT(*) AS num_orders
FROM
 `business_case.orders`
GROUP BY
 order_month
ORDER BY
 order_month;
```

Query results

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

Row	order_month	num_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

Grasping monthly seasonality facilitates proactive decision-making, guaranteeing the business is adequately equipped to handle fluctuations in customer demand year-round. These insights can be leveraged for proficient marketing, inventory management, and resource planning.

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

```
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 order_time_of_day,
 COUNT(*) AS num_orders
FROM
 `business_case.orders`
GROUP BY
 order_time_of_day;
```

Query results

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

Row	order_time_of_day	num_orders
1	Morning	27733
2	Dawn	5242
3	Afternoon	38135
4	Night	28331

The afternoon emerges as the peak period for customer transactions, boasting the highest order volume at 38,135. Following closely are the morning and night periods, which also exhibit significant order numbers, while dawn shows a comparatively lower order count. Adjusting staffing levels, logistical operations, and customer support resources to match peak order times could be beneficial. Moreover, introducing targeted promotions or flash sales during high-order times could enhance sales and customer satisfaction.

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

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

```
SELECT
 COUNT(o.order_id) AS no_of_orders,
 EXTRACT(year
 FROM
  order_purchase_timestamp) AS order_year,
 EXTRACT(month
 FROM
  order_purchase_timestamp) AS order_month,
 c.customer_state,
FROM
 `business_case.orders` AS o
JOIN
 `business_case.customers` AS c
ON
 o.customer_id = c.customer_id
GROUP BY
 order_year,
 order_month,
 c.customer_state
ORDER BY
 no_of_orders DESC,
 order_year,
 order_month,
 c.customer_state;
```

Query results

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

Row	no_of_orders	order_year	order_month	customer_state
1	113	2016	10	SP
2	56	2016	10	RJ
3	40	2016	10	MG
4	24	2016	10	RS
5	19	2016	10	PR
6	11	2016	10	SC
7	9	2016	10	GO
8	8	2016	10	CE
9	7	2016	10	PE
10	6	2016	10	DF
11	4	2016	10	BA

Based on the query above, we will extract information regarding the number of orders placed in various years and months, categorized by customer state. Here are the insights derived from the data:

- Order Distribution : The quantity of orders varies among different states, years, and months, with some states consistently exhibiting higher order volumes compared to others.
- Seasonal Trends : Fluctuations in order volumes are observed across different months and years. Certain months display increased order volumes, likely influenced by factors such as seasonal demand, promotional activities, or market trends.
- State-wise Analysis : Analyzing the data allows for understanding the ordering patterns of different states. Some states consistently maintain high order volumes throughout the year, while others may experience fluctuations. This analysis can unveil potential market opportunities or areas requiring improvement.

- Month-wise Analysis : Examining the data on a monthly basis offers insights into ordering patterns within each year. Identifying months with the highest and lowest order volumes assists in planning inventory, marketing campaigns, and resource allocation effectively.

- Historical Comparison : Comparing order volumes across different years and months enables the identification of growth and decline trends. This information proves valuable for setting targets, forecasting, and making informed decisions.

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

```
SELECT
 COUNT(o.order_id) AS no_of_orders,
 c.customer_state,
FROM
 `business_case.orders` AS o
JOIN
 `business_case.customers` AS c
ON
 o.customer_id = c.customer_id
GROUP BY
 c.customer_state
ORDER BY
 no_of_orders DESC,
 c.customer_state;
```

Query results

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

Row	no_of_orders	customer_state
1	41746	SP
2	12852	RJ
3	11635	MG
4	5466	RS
5	5045	PR
6	3637	SC
7	3380	BA
8	2140	DF
9	2033	ES
10	2020	GO
11	1652	PE

Re

This query furnishes the count of orders per customer state, yielding valuable insights from the data:

State-wise Order Volumes : The data reveals the order quantities for each customer state, showcasing variations where some states, such as Sao Paulo (SP), Rio de Janeiro (RJ), and Minas Gerais (MG), exhibit comparatively higher order volumes, while others like Amapa (AP), Roraima (RR), and Acre (AC) demonstrate lower order volumes.

Market Opportunities : Leveraging the data aids in identifying potential market opportunities. States with lower order volumes, such as AP, RR, and AC, may present avenues for market expansion or targeted marketing endeavors aimed at bolstering customer engagement.

Market Share Analysis : By comparing order volumes across states, one can gauge the market share of each state. States with higher order volumes command a larger market share, indicating a potentially robust customer base and market presence.

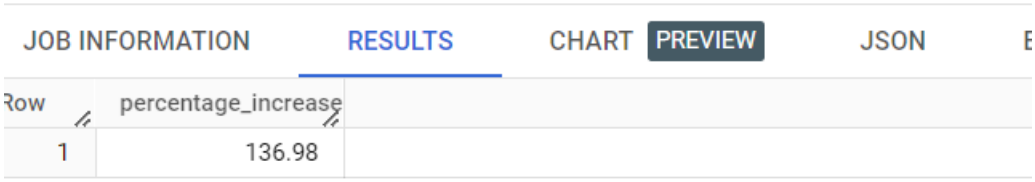
Strategic Decision-making : The data serves as a compass for strategic decision-making in businesses. It facilitates the prioritization of marketing investments, expansion plans, or customer retention strategies by emphasizing states with higher order volumes or untapped potential.

4. 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.

```
WITH
 cost_table AS (
 SELECT
  SUM(p.payment_value) AS total_cost,
  EXTRACT(year
  FROM
   order_purchase_timestamp) AS years
 FROM
  `business_case.payments` AS p
 JOIN
  `business_case.orders` AS o
 ON
  p.order_id = o.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
  years)
SELECT
 ROUND ((((
    SELECT
     total_cost
    FROM
     cost_table
    WHERE
     years = 2018) - total_cost) / total_cost * 100),2)AS percentage_increase
FROM
 cost_table
WHERE
 years = 2017;
```

Query results



The calculated percentage increase of 136.98% indicates a substantial growth in order costs from January to August between 2017 and 2018. This suggests a notable rise in the average order value or total customer expenditure over this period.

Such a high percentage increase suggests various factors contributing to this growth, including potential increases in product prices, higher customer spending, expanded product offerings, an enlarged customer base, or improved marketing strategies.

Delving into the reasons behind this significant increase can yield valuable insights into business performance and aid in making well-informed decisions for future growth. It is advisable to conduct further investigation into the underlying factors driving the growth in order costs to identify opportunities for optimizing pricing strategies, enhancing customer retention, targeting high-value customers, or implementing measures to sustain and build upon this positive trend.

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

```
SELECT
 c.customer_state,
 ROUND(SUM(oi.price), 2) AS total_order_price,
 ROUND(AVG(oi.price), 2) AS avg_order_price
FROM
 `business_case.orders` o
JOIN
 `business_case.order_items` oi
ON
 o.order_id = oi.order_id
JOIN
 `business_case.customers` c
ON
 o.customer_id = c.customer_id
GROUP BY
 c.customer_state
ORDER BY
 total_order_price DESC;
```

Query results

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

Row	customer_state	total_order_price	avg_order_price
1	SP	5202955.05	109.65
2	RJ	1824092.67	125.12
3	MG	1585308.03	120.75
4	RS	750304.02	120.34
5	PR	683083.76	119.0
6	SC	520553.34	124.65
7	BA	511349.99	134.6
8	DF	302603.94	125.77
9	GO	294591.95	126.27
10	ES	275037.31	121.91
11	PE	262788.03	145.51

Results per

Gaining insight into both total and average order values per state offers valuable information for targeted marketing, strategic planning, and resource allocation based on state-specific data. Tailoring strategies according to observed patterns in each state can be considered to maximize effectiveness and efficiency.

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

```
SELECT
 c.customer_state,
 ROUND(SUM(oi.freight_value), 2) AS total_freight_value,
 ROUND(AVG(oi.freight_value), 2) AS avg_freight_value
FROM
 `business_case.orders` o
JOIN
 `business_case.order_items` oi
ON
 o.order_id = oi.order_id
JOIN
 `business_case.customers` c
ON
 o.customer_id = c.customer_id
GROUP BY
 c.customer_state
ORDER BY
 total_freight_value DESC;
```

Query results

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

Row	customer_state	total_freight_value	avg_freight_value
1	SP	718723.07	15.15
2	RJ	305589.31	20.96
3	MG	270853.46	20.63
4	RS	135522.74	21.74
5	PR	117851.68	20.53
6	BA	100156.68	26.36
7	SC	89660.26	21.47
8	PE	59449.66	32.92
9	GO	53114.98	22.77
10	DF	50625.5	21.04
11	ES	49764.6	22.06

São Paulo (SP) leads with a total freight value of $718,723.07, with Rio de Janeiro (RJ) and Minas Gerais (MG) close behind. Analyzing total and average order freight values by state provides insights for improving shipping strategies and customer satisfaction. Tailoring promotions and strategies to regional shipping patterns can enhance efficiency and meet customer expectations.

5. 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_delivered_customer_date - order_estimated_delivery_date
-

```
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
`business_case.orders`;
```

Query results

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

Row	order_id	time_to_deliver	diff_estimated_delive
1	1950d777989f6a877539f5379...	30	12
2	2c45c33d2f9cb8ff8b1c86cc28...	30	-28
3	65d1e226dfaeb8cdc42f66542...	35	-16
4	635c894d068ac37e6e03dc54e...	30	-1
5	3b97562c3aee8bdedcb5c2e45...	32	0
6	68f47f50f04c4cb6774570cfde...	29	-1
7	276e9ec344d3bf029ff83a161c...	43	4
8	54e1a3c2b97fb0809da548a59...	40	4
9	fd04fa4105ee8045f6a0139ca5...	37	1
10	302bb8109d097a9fc6e9cefc5...	33	5
11	66057d37308e787052a32828	38	6

Result

Examining delivery time and the variance between estimated and actual delivery dates yields valuable insights regarding delivery process efficiency, customer satisfaction levels, and potential areas for operational enhancement. Addressing outliers and refining processes are crucial steps towards optimizing delivery performance

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

```
SELECT
 high.customer_state AS high_state,
 high.average_freight_value AS high_avg_freight,
 low.customer_state AS low_state,
 low.average_freight_value AS low_avg_freight
FROM (
 SELECT
  c.customer_state,
  ROUND(AVG(p.freight_value), 2) AS average_freight_value,
  ROW_NUMBER() OVER(ORDER BY  (ROUND(AVG(p.freight_value), 2)) DESC) AS rowval1
  FROM `business_case.orders` AS o
 JOIN
  `business_case.order_items` AS p
 ON
  o.order_id = p.order_id
 JOIN
  `business_case.customers` AS c
  ON o.customer_id= c.customer_id
 GROUP BY
  c.customer_state
 ORDER BY
  average_freight_value DESC
 LIMIT
  5 ) AS high
JOIN (
 SELECT
  c.customer_state,
  ROUND(AVG(p.freight_value), 2) AS average_freight_value,
  ROW_NUMBER() OVER(ORDER BY (ROUND(AVG(p.freight_value), 2)))
  AS rowval2
 FROM
  `business_case.orders` AS o
 JOIN
  `business_case.order_items` AS p
 ON
  o.order_id = p.order_id
 JOIN
```

```
  `business_case.customers` AS c
 ON
  o.customer_id = c.customer_id
 GROUP BY
  c.customer_state
 ORDER BY
  average_freight_value
 LIMIT
  5 ) AS low
ON
 high.rowval1 = low.rowval2;
```

Query results

SAVE R

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

Row	high_state	high_avg_freight	low_state	low_avg_freight
1	RR	42.98	SP	15.15
2	PB	42.72	PR	20.53
3	RO	41.07	MG	20.63
4	AC	40.07	RJ	20.96
5	PI	39.15	DF	21.04

This query provides information regarding the average freight value by customer state. Here are insights and recommendations:

Average Freight Value : The "avg_freight_value" column denotes the average freight cost associated with each customer state. This metric represents the average expense incurred for shipping goods to customers within a specific state. Analyzing this data helps identify disparities in freight costs across various regions and states.

Freight Cost Comparison : By comparing the average freight values of different states, it's evident that RR (Roraima) has the highest freight value at 42.98, followed by PB (Paraiba) at 42.72, RO (Rondonia) at 41.07, AC (Acre) at 40.07, and PI (Piaui) at 35.15. This suggests that shipping goods to RR generally entails higher freight costs compared to other states in the dataset.

Lower freight costs in other states might be attributed to factors such as proximity to shipping hubs, well-established transportation infrastructure, or a competitive logistic market.

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

```
WITH
 cte AS(
SELECT
 c.customer_state,
 ROUND (AVG(t1.delivery_time), 2) AS avg_delivery_time
FROM (
 SELECT *,
  TIMESTAMP_DIFF(order_delivered_customer_date,
   order_purchase_timestamp,
   day) AS delivery_time,
 FROM
  `business_case.orders`
 WHERE
  order_status= 'delivered'
  AND order_delivered_customer_date IS NOT NULL
 ORDER BY
   order_purchase_timestamp ) AS t1
JOIN
 `business_case.customers` AS c
ON
 t1.customer_id = c.customer_id
GROUP BY
 c.customer_state
ORDER BY
 avg_delivery_time )
SELECT
 c1.customer_state AS low_state,
 c1.avg_delivery_time AS low_avg_delivery_time,
 c2.customer_state AS high_state,
 c2.avg_delivery_time AS high_avg_delivery_time
FROM (
 SELECT
  *, ROW_NUMBER() OVER (ORDER BY cte.avg_delivery_time DESC) AS rowval2
 FROM
  cte
 ORDER BY
  rowval2) AS c2
 JOIN
 (
 SELECT
 *,
  ROW_NUMBER() OVER (ORDER BY cte.avg_delivery_time) AS rowval1
 FROM cte
 ORDER BY
  rowval1 ) AS c1
ON
 c1.rowval1= c2.rowval2
LIMIT
 5;
```

Query results

SAVE

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

Row	low_state	low_avg_delivery_tim	high_state	high_avg_delivery_tir
1	SP	8.3	RR	28.98
2	PR	11.53	AP	26.73
3	MG	11.54	AM	25.99
4	DF	12.51	AL	24.04
5	SC	14.48	PA	23.32

- Identifying regions with efficient delivery operations, shorter transit times, or robust logistics networks involves examining states like SP and PR, which exhibit the lowest average delivery times, and contrasting them with states like RR and AP, which have the highest average delivery times.
- These insights prove valuable for our company's efforts to enhance customer satisfaction, optimize operational efficiency in delivery processes, and establish realistic expectations for customers based on regional delivery time trends.
- While analyzing the data and deriving conclusions from these insights, it's essential to consider additional factors such as population density, the distinction between urban and rural areas, customer expectations, and unique logistical constraints.
- Armed with this information, our company can focus on areas where improvements in delivery efficiency are possible, ultimately enhancing customer experiences and operational effectiveness.

D. Find out the top 5 states where the order delivery is really fast 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.

```
SELECT
 c.customer_state,
 ROUND(AVG(DATE_DIFF(o.order_estimated_delivery_date,o.order_delivered_customer_date,day)),2) AS
delivery_diff,
FROM
 `business_case.orders` AS o
JOIN
 `business_case.customers` AS c
ON
 c.customer_id = o.customer_id
GROUP BY
 c.customer_state
ORDER BY
 delivery_diff ASC
LIMIT
 5;
```

Query results

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

Row	customer_state	delivery_diff
1	AL	7.95
2	MA	8.77
3	SE	9.17
4	ES	9.62
5	BA	9.93

Job history

States with faster delivery: those with the lowest average delivery difference are deemed to have swift delivery compared to the estimated date. Recognizing these states for their efficient delivery processes is essential, as customers in these regions are likely to have received their orders promptly, contributing to a positive customer experience.

6. Analysis based on the payments

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

```
select count(o.order_id) no_of_order,p.payment_type,
extract(month from order_purchase_timestamp)as months,
extract(year from order_purchase_timestamp) as years
from `business_case.orders` as o join `business_case.payments` as p
on o.order_id = p.order_id
group by months,years,p.payment_type
order by years,months
```

Query results

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

Row	no_of_order	payment_type	months	years
1	3	credit_card	9	2016
2	254	credit_card	10	2016
3	63	UPI	10	2016
4	23	voucher	10	2016
5	2	debit_card	10	2016
6	1	credit_card	12	2016
7	583	credit_card	1	2017
8	197	UPI	1	2017
9	61	voucher	1	2017
10	9	debit_card	1	2017
11	1356	credit_card	2	2017

Result

Job history

The comparison underscores the enduring popularity of specific payment methods, offering insights for strategic planning and tailored promotional activities to meet customer preferences across different months. Consider implementing the following steps:

- Concentrate on marketing strategies and promotions that resonate with the popularity of UPI and credit cards.
- Customize campaigns to increase voucher usage, possibly by offering incentives to encourage adoption.
- Monitor and incentivize debit card usage through potential targeted promotions.

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

```
SELECT
 payment_installments,
 COUNT(*) AS num_orders
FROM
 `business_case.payments`
GROUP BY
 payment_installments
ORDER BY
 payment_installments;
```

Query results

JOB INFORMATION | RESULTS | CHART | PREVIEW | J

Row	payment_installment	num_orders
1	0	2
2	1	52546
3	2	12413
4	3	10461
5	4	7098
6	5	5239
7	6	3920

Analysis of payment installments in orders reveals a preference for single-payment transactions, but also interest in installment plans. Most customers opt for 1, 2, 3, or 4 installments, with decreasing frequency for higher numbers. Rarely do orders exceed 10 payments. Strategic considerations should target promotions for higher installment plans, understand customer behavior affecting preferences, and evaluate installment impact on revenue. Tailoring marketing strategies accordingly will enhance the shopping experience and optimize payment plans.

Actionable Insights :

Based on the given dataset , the following actionable insights can be drawn.

- Identify the peak time : Brazilian clients frequently order more in the afternoons, indicating that this is a period when people want to shop online. By scheduling customer assistance personnel or launching focused marketing efforts during the busiest ordering periods, for example, we may operate more efficient operations with the aid of this information. Also, customers are buying least during dawn

- Identify the Peak Periods : In 2016, there were 329 orders, indicating the initial year of data collection. The subsequent years show substantial growth with 2017 witnessing a significant increase to 45,101 orders and 2018 continued the upward trend with 54,011 orders. We can use this information for forecasting future order volumes and planning resources accordingly

- State-wise Analysis : it is observed that the state of São Paulo (SP) had the highest count of customers with 41,74 , while RR had the least with only 46 customers. This information can be used to focus on high- performing states for marketing and sales efforts and identify potential opportunities in states with low customer counts.

- Analyze Freight Values and Delivery Time : The top 5 states with the highest average freight values & the states with the highest average time to delivery information can help optimize freight costs and delivery times by identifying areas that need improvement or potential cost-saving opportunities in cases like RR- the state which has highest avg fright value and also have less customers

- Identify Delays and Issues : With the given data in some places the maximum time difference between purchase and delivery was approximately 188 days, while the estimated delivery time was 19 days. This indicates potential delays and issues in the delivery process that need to be addressed to improve customer satisfaction and optimize delivery operations. Additionally, state RR have maximum average delivery time with least customers.

- Payment Modes and Installments : The majority of orders were placed through credit card and UPI payment modes. Additionally, the count of orders was highest for 1 installment and lowest for 23 installments. This information can be used to tailor payment options and installment plans to meet customer preferences and optimize sales.

Recommendations :

- To enhance customer satisfaction and optimize resource allocation, it's imperative to pinpoint peak seasons and adjust marketing strategies accordingly.
- Utilizing insights into monthly seasonality can inform inventory management, marketing initiatives, and resource distribution.
- Monitoring trends in cost escalation facilitates informed decisions regarding pricing strategies and cost management.
- Targeted promotions in states with high order values can yield higher returns, while exploring opportunities in states with lower order values may expand market share.
- Improving freight and delivery processes is essential for enhancing overall customer satisfaction, especially in states with high average delivery times and freight costs.
- Analyzing the popularity of payment methods and understanding customer preferences for payment installments allows for optimization of payment gateways and targeted promotional efforts.

Implementing these actionable recommendations collectively fosters a more efficient, customer-centric, and economically optimized e-commerce operation.