TARGET's Brazil E-com Market EDA on Google Big Query

Analysis By: Suman Kumar Nandi

Context: Target is one of the world's most recognized brands and one of America's leading retailers. Target makes itself a preferred shopping destination by offering outstanding value, inspiration, innovation, and an exceptional guest experience that no other retailer can deliver.
This business case has information on 100k orders from 2016 to 2018 made at Target in Brazil. Its features allow viewing orders from multiple dimensions: from order status, price, payment, and freight performance to customer location, product attributes, and finally reviews written by customers.

1. Import the dataset and do the usual exploratory analysis steps like checking the structure & characteristics of the dataset
1.1Data type of columns in a table
1.2Time period for which the data is given
1.3Cities and States of customers ordered during the given period
Data Description:

we have a total of 8 tables in the form of a dataset.
1. customers.csv
2. geolocation.csv
3. order_items.csv
4. payments.csv
5. reviews.csv
6. orders.csv
7. products.csv
8. Sellers.csv

1.1 Data type of columns in a table:

Customers Table:

Field name	Type
customer_id	STRING
customer_unique_id	STRING
customer_zip_code_prefix	INTEGER
customer_city	STRING
customer_state	STRING

Geolocation Table:

TARGET's Brazil E-com Market EDA on Google Big Query

Analysis By: Suman Kumar Nandi

Field name	Type
geolocation_zip_code_prefix	INTEGER
geolocation_lat	FLOAT
geolocation_lng	FLOAT
geolocation_city	STRING
geolocation_state	STRING

Order_Item Table:

Field name	Type
order_id	STRING
order_item_id	INTEGER
product_id	STRING
seller_id	STRING
shipping_limit_date	TIMESTAMP
price	FLOAT
freight_value	FLOAT

Payments Table:

Field name	Type
order_id	STRING
payment_sequential	INTEGER
payment_type	STRING
payment_installments	INTEGER
payment_value	FLOAT

Reviews Table:

Field name	Type
review_id	STRING
order_id	STRING
review_score	INTEGER
review_comment_title	STRING
review_creation_date	TIMESTAMP
review_answer_timestamp	TIMESTAMP

TARGET's Brazil E-com Market EDA on Google Big Query

Analysis By: Suman Kumar Nandi

Orders Table:

Field name	Type
order_id	STRING
customer_id	STRING
order_status	STRING
order_purchase_timestamp	TIMESTAMP
order_approved_at	TIMESTAMP
order_delivered_carrier_date	TIMESTAMP
order_delivered_customer_date	TIMESTAMP
order_estimated_delivery_date	TIMESTAMP

Products Table:

Field name	Type
product_id	STRING
product_category	STRING
product_name_length	INTEGER
product_description_length	INTEGER
product_photos_qty	INTEGER
product_weight_g	INTEGER
product_length_cm	INTEGER
product_height_cm	INTEGER

Sellers Table:

Field name	Type
seller_id	STRING
seller_zip_code_prefix	INTEGER
seller_city	STRING
seller_state	STRING

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

```
SELECT
 MIN(order_purchase_timestamp) AS MIN_TIME,
 MAX(order_purchase_timestamp) AS MAX_TIME
FROM `target-sql-423910.target_sql.orders`
```

TARGET's Brazil E-com Market EDA on Google Big Query

Analysis By: Suman Kumar Nandi

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

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

```
select
   count(distinct customers.customer_city) as count_city,
   count(distinct customers.customer_state) as count_state
from `target_sql.customer_purchases` as customers
join `target_sql.orders` as ord
on customers.customer_id = ord.customer_id
```

Row	count_city	count_state
1	4119	27

2. In-depth Exploration:

1.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 order_count
FROM `target_sql.orders`
where order_delivered_carrier_date is not null
GROUP BY
  order_year
ORDER BY
  order_year;
```

Row	order_year	order_count
1	2016	281
2	2017	43957
3	2018	53420

With the above query, we will get the total no of orders with respect to the year so in the year 2016 total no of orders was 281, in 2017 it become 43957 orders, and in 2018 this number was 53420. So from here, we can see year-on-year growth in the no of orders.

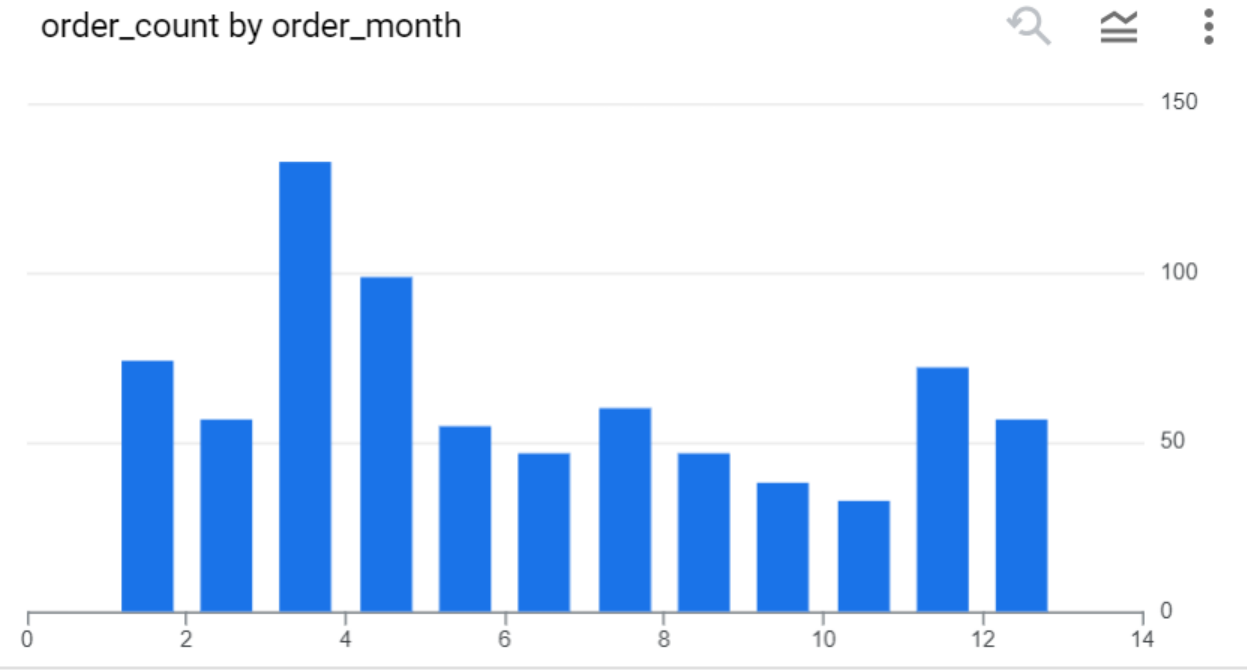
2.Can we see some kind of monthly seasonality in terms of the no. of orders being placed?

TARGET's Brazil E-com Market EDA on Google Big Query

Analysis By: Suman Kumar Nandi

```
SELECT
  EXTRACT(MONTH FROM order_purchase_timestamp) AS order_month,
  COUNT(*) AS order_count
FROM
  `target-sql-423910.target_sql.orders`
where lower(order_status) LIKE "%shipped%"
GROUP BY
   order_month
ORDER BY
   order_month;
```

With the above query, we will get orders w.r.t months that from Mid of November till Mid of April there is a increase in average order counts and in March and April, the sales is at peak.
So we can say that in Brazil Maximum Sales order is during the March and April of Every Year.



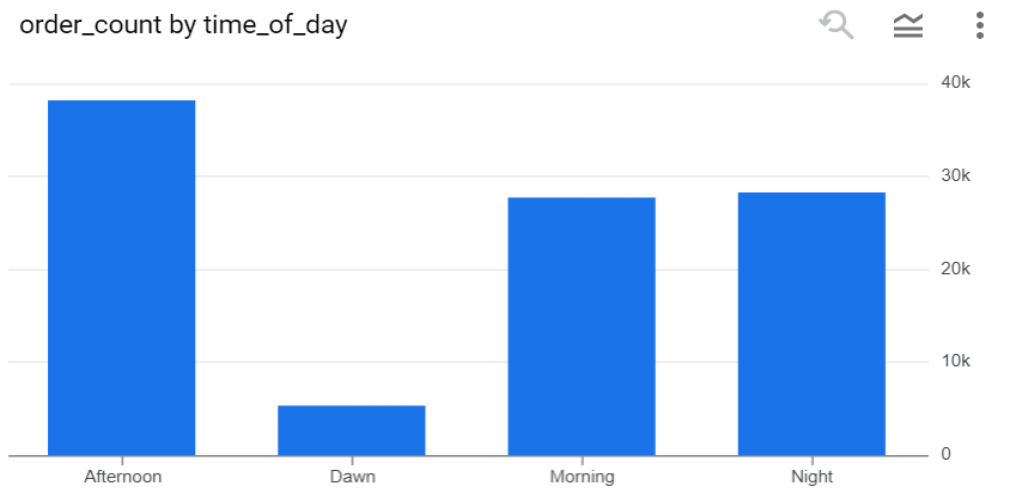
3. During what time of the day, do the Brazilian customers mostly place their orders? (Dawn, Morning, Afternoon or Night)

TARGET's Brazil E-com Market EDA on Google Big Query

Analysis By: Suman Kumar Nandi

1.
   - 0-6 hrs : Dawn
   - 7-12 hrs : Mornings
   - 13-18 hrs : Afternoon
   - 19-23 hrs : Night

```
WITH OrderTimes AS (
  SELECT
    DISTINCT customer_id,
    EXTRACT(HOUR FROM order_purchase_timestamp) AS order_hour
  FROM
    `target-sql-423910.target_sql.orders`
)
SELECT
  CASE
    WHEN order_hour >= 0 AND order_hour < 7 THEN 'Dawn'
    WHEN order_hour >= 7 AND order_hour < 13 THEN 'Morning'
    WHEN order_hour >= 13 AND order_hour < 19 THEN 'Afternoon'
    ELSE 'Night'
  END AS time_of_day,
  COUNT(*) AS order_count
FROM
  OrderTimes
GROUP BY
  time_of_day
ORDER BY
  time_of_day;
```



From the above table, we can see that Brazilian customers tend to buy more in the afternoon.

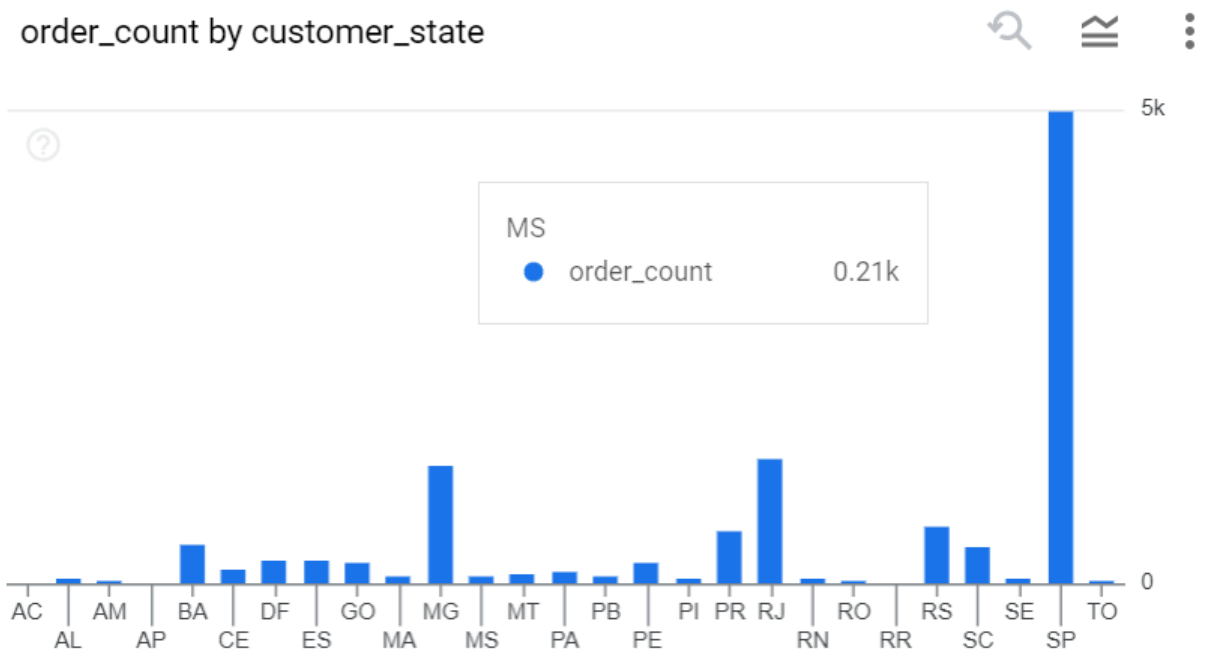
3. Evolution of E-commerce orders in the Brazil region:

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

TARGET's Brazil E-com Market EDA on Google Big Query

Analysis By: Suman Kumar Nandi

```sql
SELECT
    c.customer_state,
    EXTRACT(MONTH FROM o.order_purchase_timestamp) AS order_month,
    COUNT(*) AS order_count
FROM
    `target-sql-423910.target_sql.orders` AS o
JOIN
    `target_sql.customer_purchases` AS c
ON
    o.customer_id = c.customer_id
GROUP BY
    c.customer_state, order_month
ORDER BY
     c.customer_state, order_month;
```



From the above query, we will get the information about the number of orders placed in different years and months, categorized by customer state. Here are some insights from the data.

1. Distribution of orders: the number of orders varies across different states, years, and months. some states have higher no of orders compared to others.

2. Seasonality: there is a fluctuation in the number of orders over different months and years. Certain months show higher order volumes. This could be attributed to various factors such as seasonal demand, promotions, or market trends.

3.state_wise Analysis: you can analyze the data to understand the ordering

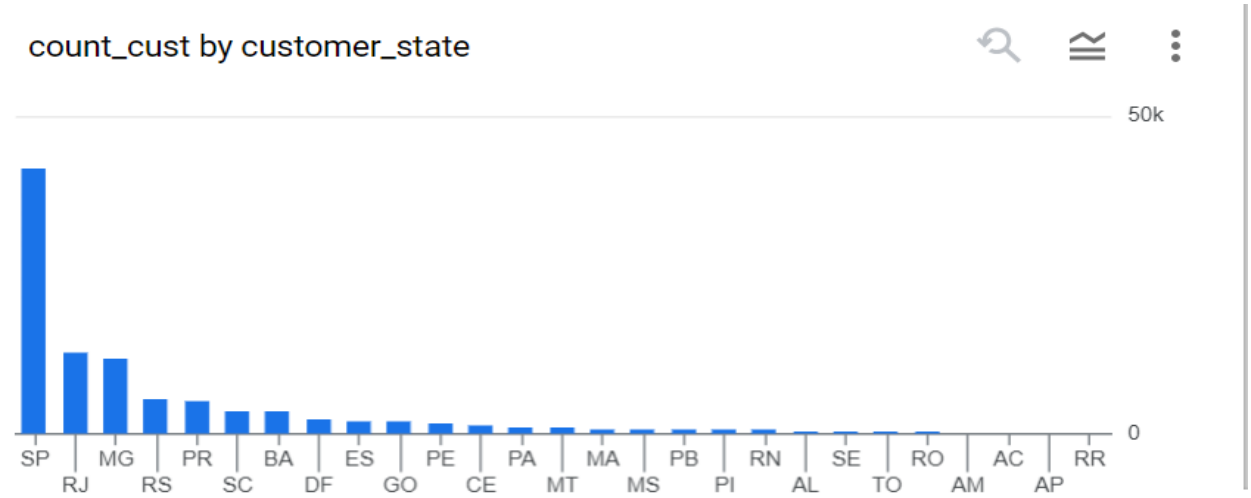
pattern of different states. Some states have constantly had a high number of orders throughout year, while others may have fluctuations. This information can be used to identify potential market opportunities or areas of improvement.

4.month_wise analysis: Analyzing the data on a monthly basis provides insights into the ordering pattern within each year. You can identify which month have the highest and lowest order volumes, helping to plan your inventory, marketing campaign, and resource allocation accordingly.

5. Historical Comparison: by comparing the order volumes across different years and months, we can identify any growth and decline trends. This information can be valuable for setting targets, forecasting, and making informed decisions.

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

```
SELECT
  count(customer_id) as count_cust,
  customer_state
from `target_sql.customer_purchases`
group by customer_state
order by  count_cust desc
```



This query will provide you the number of orders for each customer state. Here are some insights derived from the data.

1.state_wise order volumes: The data indicates the number of orders for each customer state. Some states have a higher number of orders compared to others. like Sao Paula(SP), Rio de Janeiro(RJ), and Minas Gerais(MG) have a relatively higher order volumes, while states like Amapa(AP), Roraima(RR), and Acre(AC)

have lower order volumes.

2. Market Opportunities: the data can help identify potential market opportunities. States with relatively lower order volumes, such as AP, RR, and AC, may present opportunities for market expansion or targeted marketing efforts to increase customer engagement.

3. Market Share: Comparing the order volumes of different states allows you to assess the market share of each state. States with higher-order columns have a larger market share. Indicating a potentially stronger customer base and market presence.

4. Strategic Decision-making: the data guide strategic decision-making for businesses. It can help prioritize marketing investments, expansion plans, or customer retention strategies by focusing on 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.

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

```sql
WITH TotalCostByYear AS (
  SELECT
    EXTRACT(YEAR FROM order_purchase_timestamp) AS order_year,
    SUM(price) AS total_cost
  FROM
    `target-sql-423910.target_sql.order_items` as od_item
  join `target_sql.orders` as od on od_item.order_id = od.order_id
  WHERE
    EXTRACT(MONTH FROM order_purchase_timestamp) BETWEEN 1 AND 8
    AND EXTRACT(YEAR FROM order_purchase_timestamp) IN (2017, 2018)
    AND order_status like "%delivered%"
  GROUP BY
    order_year)
SELECT
  t1.order_year,
  t1.total_cost AS total_cost,
  t2.total_cost AS previous_year_total_cost,
  ((t1.total_cost - t2.total_cost) / t2.total_cost) * 100 AS percentage_increase
FROM
  TotalCostByYear t1
JOIN
  TotalCostByYear t2
ON t1.order_year = t2.order_year + 1
WHERE
  t1.order_year = 2018;
```

TARGET's Brazil E-com Market EDA on Google Big Query

Analysis By: Suman Kumar Nandi

order_year	total_cost	previous_year_total_cost	percentage_increase
2018	7218125.0	2993456.0	141.0

The calculated percentage increase of 141% suggests a significant growth in order cost from 2017 to 2018 between January to August. This indicates that the avg order value or the total expenditure by customers has substantially increased over the specified time period.

The high percentage increase implies that there might be factors contributing to this growth, such as an increase in product prices, higher customer spending, expanded product offerings, increased customer base, or improved marketing strategies.

Analyzing the reasons behind this significant increase can provide valuable insights into the business performance and help make informed decisions for future growth. It is recommended to further investigate the underlying factors contributing to the growth in order cost to identify opportunities for optimizing pricing strategies,

It is enhancing customer retention, targeting high-value customers, or implementing measures to sustain and build upon the positive trend.

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

```
select
sum(price) as total_value,
avg(price) as avg_value,
cp.customer_state
from `target_sql.order_items` as odt
join `target_sql.orders` as od on odt.order_id = od.order_id
join `target_sql.customer_purchases` as cp on od.customer_id = cp.customer_id
group by cp.customer_state
order by total_value desc, avg_value desc
```

total_value	avg_value	customer_state
5202955.05	109.65	SP
1824092.67	125.12	RJ
1585308.03	120.75	MG
750304.02	120.34	RS
683083.76	119.0	PR
520553.34	124.65	SC
511349.99	134.6	BA
302603.94	125.77	DF
294591.95	126.27	GO
275037.31	121.91	ES

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

TARGET's Brazil E-com Market EDA on Google Big Query

Analysis By: Suman Kumar Nandi

```
select
round(sum(freight_value),2) as total_value,
round(avg(freight_value),2) as avg_value,
cp.customer_state
from `target_sql.order_items` as odt
join `target_sql.orders` as od on odt.order_id = od.order_id
join `target_sql.customer_purchases` as cp on od.customer_id = cp.customer_id
group by cp.customer_state
order by total_value desc, avg_value desc
```

total_value	avg_value	customer_state
718723.07	15.15	SP
305589.31	20.96	RJ
270853.46	20.63	MG
135522.74	21.74	RS
117851.68	20.53	PR
100156.68	26.36	BA
89660.26	21.47	SC
59449.66	32.92	PE
53114.98	22.77	GO
50625.5	21.04	DF

From the above query provides us with various metrics such as the total sum of the price, the total sum
of freight, the average price, average freight value, total sum, total avg, and customer state. Here are
some insights and recommendations.

1. Revenue and freight Analysis: The total sum of price represents the overall revenue generated
from orders, while the total sum of freight indicates the total shipping cost. By analyzing these we can assess the revenue and freight expenditure for each state. Like Sao Paulo(SP), Rio de Janeiro(RJ), and Minas Gerais(MG) have higher revenue and freight costs, suggesting a large market and potentially higher shipping requirements.

2. Avg price and freight: The avg price and avg freight value provide insights into the pricing and
shipping patterns across the state. State with higher avg prices may indicate a preference for more expensive products or a willingness to spend more, similarly higher avg freight value may
suggest longer shipping distances or higher shipping costs. Understanding these patterns can help businesses tailor pricing strategies and optimize shipping options.

3. Pricing and Shipping Strategies: Analyzing the average price and freight values can guide pricing and shipping strategies. State with higher avg prices may allow you for premium pricing
strategies, while state with higher average freight values may require efficient shipping solutions or potential partnerships with logistic providers to optimize cost.

4. Customer Behavior: the data from this query provides insights into customer behavior within
different states. A state with a higher average price may indicate a market segment that values
quality or is willing to spend more on certain products. This information can be leveraged for product assortment, marketing messaging, and customer segmentation strategies.

5. Competitive Analysis: Consider benchmarking the metrics against competitors or industry averages to assess your business performance. Understanding how your metrics compare to the market can provide insights into competitiveness and areas for improvement

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

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

```
SELECT
  order_id,
  DATE_DIFF(order_delivered_customer_date,order_purchase_timestamp, DAY) AS time_of_deliver,
  DATE_DIFF(order_estimated_delivery_date,order_purchase_timestamp, DAY) AS
estimate_delivery_time,
  DATE_DIFF(order_estimated_delivery_date,order_purchase_timestamp, DAY) -
  DATE_DIFF(order_delivered_customer_date,order_purchase_timestamp, DAY) AS
delivery_diff_estimated_days
FROM
  `target-sql-423910.target_sql.orders`
WHERE DATE_DIFF(order_delivered_customer_date,order_purchase_timestamp, DAY) is not null
```

TARGET's Brazil E-com Market EDA on Google Big Query

Analysis By: Suman Kumar Nandi

Row	order_id	time_of_deliver	estimate_delivery_tir	delivery_diff_estimat
1	1950d777989f6a877539f5379...	30	17	-13
2	2c45c33d2f9cb8ff8b1c86cc28...	30	59	29
3	65d1e226dfaeb8cdc42f66542...	35	52	17
4	635c894d068ac37e6e03dc54e...	30	32	2
5	3b97562c3aee8bdedcb5c2e45...	32	33	1
6	68f47f50f04c4cb6774570cfde...	29	31	2
7	276e9ec344d3bf029ff83a161c...	43	39	-4
8	54e1a3c2b97fb0809da548a59...	40	36	-4
9	fd04fa4105ee8045f6a0139ca5...	37	35	-2

This query provides you information about order delivery, estimated delivery time, and the difference between the actual delivery time and the estimated delivery time. Here are some insight and recommendations.

1. Delivery Performance: Analyzaing the time or delivery and the estimated delivery time can help asses the performance of the delivery process. The "Diff_time of_delivery_estimate_delivery" column represent the diff between the actual delivery time and the estimated delivery time. Negative values indicate that the order was delivered earlier than estimate, while positive values indicate a delay in delivery.

2. Delivery Efficiency: Reviewing the difference between the time of delivery and the estimated delivery time can help identify patterns or issues in the delivery process. Negative values consistently across multiple orders may indicate an efficient delivery system, while positive values may suggest the need to improve delivery logistics or manage customer expectations regarding delivery times.

3. Customer Satisfaction: Timely delivery is an essential factor in customer satisfaction. Monitoring the difference between the actual and estimated delivery times can help identify potential areas of improvement to enhance customer experience.

4. Operational Optimization: we can identify areas of operational improvement. For instance, if there are consistent delay(positive values) in the delivery process, it may be necessary to review and optimize logistics operations, inventory management, order processing, and transportation methods
5. Continuous Monitoring: Regularly monitoring delivery performance metrics and customer feedback is essential for ongoing improvement. By tracking the time of delivery, estimated delivery time, and areas that require attention. Implementing a system to monitor delivery performance can help address issues promptly and improve overall efficiency.

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

TARGET's Brazil E-com Market EDA on Google Big Query

Analysis By: Suman Kumar Nandi

```
----highest average freight value


WITH AvgFreightByState AS (
  SELECT
    customer_state,
    ROUND(AVG(freight_value),2) AS avg_freight_value
  FROM
    `target-sql-423910.target_sql.customer_purchases` as cp
  JOIN `target_sql.orders` as o on cp.customer_id = o.customer_id
  JOIN `target_sql.order_items` as ot on o.order_id = ot.order_id
  GROUP BY
    customer_state
)
SELECT
  customer_state,
  avg_freight_value
FROM
  AvgFreightByState
ORDER BY
  avg_freight_value DESC
LIMIT 5
```

Row	avg_freight_value	customer_state
1	42.98	RR
2	42.72	PB
3	41.07	RO
4	40.07	AC
5	39.15	PI

This query provide us information about the average freight value and customer state, here are some insight and recommendation.

1. Average Freight Value: the "avg_freight_value" column indicates the avg freight value associated with each customer state. This value represents the avg cost of shipping goods to customers in that particular state. Analyzing this data helps identify variations in freight cost across different region and states.
2. Freight cost comparison: Comparing the avg freight values of different states, we can observe
that RR(Roraiba) 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 indicates that shipping
goods to RR generally incur higher freight costs compared to another state in the dataset.

```
----lowest average freight value
```

TARGET's Brazil E-com Market EDA on Google Big Query

Analysis By: Suman Kumar Nandi

```
WITH AvgFreightByState AS (
  SELECT
    customer_state,
    ROUND(AVG(freight_value),2) AS avg_freight_value
  FROM
    `target-sql-423910.target_sql.customer_purchases` as cp
  JOIN `target_sql.orders` as o on cp.customer_id = o.customer_id
  JOIN `target_sql.order_items` as ot on o.order_id = ot.order_id
  GROUP BY
    customer_state
)
SELECT
  customer_state,
  avg_freight_value
FROM
  AvgFreightByState
ORDER BY
  avg_freight_value ASC
LIMIT 5
```

Row	avg_freight_value	customer_state
1	15.15	SP
2	20.53	PR
3	20.63	MG
4	20.96	RJ
5	21.04	DF

This query provide us data thee avg freight value and customer state, here are the recommendation top 5 states with the lowest average freight value.

1. SP(Sao Paulo) - Average freight value: 15.15
2. PR(Parana) - Average freight value: 20.53
3. MG(Minas Gerais) - Average freight value: 20.63
4. RJ(Rio de Janeiro) - Average freight value: 20.96
5. DF (Distrito Federal) - Average freight value: 21.04

1. These states have relatively lower avg fright value compared to other states.

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

3. Identify Best Practices: identify the logistics practices and strategies implemented in these states that contribute to the lower freight cost. Identify any best practices, operational efficiencies, or collaboration

with logistics partners that can be adopted in other regions to optimize freight expenses.

4. Suppliers networks: explore building or expanding supplier networks in these states.

5. Negotiate Shipping rates: leverage the lower avg freight value as bargaining power when negotiating shipping rates with logistics partners.

6. Consolidate shipments: maximize container or vehicle utilization, by a combination of multiple orders in single shipments.

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

```
----highest average delivery times


WITH
  DeliveryTimeByState AS (
  SELECT
    customer_state,
    ROUND(AVG(TIMESTAMP_DIFF(order_delivered_customer_date, order_purchase_timestamp, DAY)))
AS avg_delivery_time
  FROM `target_sql.orders` as o
  JOIN `target_sql.customer_purchases` as cp
  on o.customer_id = cp.customer_id
  GROUP BY
    customer_state )

SELECT
  customer_state,
  avg_delivery_time
FROM
  DeliveryTimeByState
ORDER BY
  avg_delivery_time DESC
LIMIT 5
```

Row	customer_state	avg_delivery_time
1	RR	29.0
2	AP	27.0
3	AM	26.0
4	AL	24.0
5	PA	23.0

1. These states have relatively higher avg time to delivery compared to other states in the dataset.

TARGET's Brazil E-com Market EDA on Google Big Query

Analysis By: Suman Kumar Nandi

2. Factors contributing to longer delivery time could include geographical challenges, limited transportation infrastructure, or logistical complexities in reaching these regions.
3. Logistics network optimization:
4. Carrier Performance Monitoring
5. Inventory management
6. Last-mile delivery strategies
7. Proactive communication
8. Customer feedback analysis
9. Collaboration with local partner

----lowest average delivery times

```
WITH
  DeliveryTimeByState AS (
  SELECT
    customer_state,
    ROUND(AVG(TIMESTAMP_DIFF(order_delivered_customer_date, order_purchase_timestamp, DAY)))
AS avg_delivery_time
  FROM `target_sql.orders` as o
  JOIN `target_sql.customer_purchases` as cp
  on o.customer_id = cp.customer_id
  GROUP BY
    customer_state )

SELECT
  customer_state,
  avg_delivery_time
FROM
  DeliveryTimeByState
ORDER BY
  avg_delivery_time ASC
LIMIT
  5;
```

Row	customer_state	avg_delivery_time
1	SP	8.0
2	MG	12.0
3	PR	12.0
4	DF	13.0
5	SC	14.0

1. These states have relatively lower time to delivery compared to other states.
2. Factors contributing to faster delivery times could develop well-developed transportation infrastructure, proximity to distribution centers, and efficient logistics operations in these regions

3. Maintain high-level services
4. Scalability Planning
5. Operational efficiency
6. Collaboration with carriers
7. Data-driven insights
8. Proactive issue resolution

4. Find out the top 5 states where the order delivery is really fast as compared to the estimated date of delivery.

```
WITH
  DeliveryTimeByState AS (
  SELECT
    customer_state,
    ROUND(AVG(TIMESTAMP_DIFF(order_estimated_delivery_date,order_delivered_customer_date,DAY
))) AS  avg_early_delivery
  FROM `target_sql.orders` as o
  JOIN `target_sql.customer_purchases` as cp
  on o.customer_id = cp.customer_id
  GROUP BY
    customer_state )

SELECT
  customer_state,
  avg_early_delivery
FROM
  DeliveryTimeByState
ORDER BY
  avg_early_delivery ASC
LIMIT 5
```

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

1. This query provides us with the avg delivery difference between the estimated delivery date and the actual delivery date for each customer state, here are some insight and recommendations:

2. State with the faster delivery: the state with the lowest avg delivery difference are considered to have fast delivery compared to the estimated date. These states should be recognized for their efficient delivery process customers in these states likely have received their orders quickly, which can be considered a positive customer experience

3. States with slow delivery: the state with the highest average delivery indicates slower delivery compared to the estimated date. This could be due to various factors such as logistics challenges, remote locations, or operational inefficiencies. It's important to investigate the reason behind the slow delivery and take appropriate action.

TARGET's Brazil E-com Market EDA on Google Big Query

Analysis By: Suman Kumar Nandi

Recommendations:

1. Identify bottlenecks: identify the delivery process in states to identify potential bottlenecks. Look for areas where delays commonly occur.
2. Optimize Logistics
3. Communicate Estimate Delivery date Accurately
4. Monitor performance
5. Customer feedback and satisfaction

6. Analysis based on the payments:

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

```
SELECT
  EXTRACT(YEAR FROM order_purchase_timestamp) AS order_year,
  EXTRACT(MONTH FROM order_purchase_timestamp) AS order_month,
  payment_type,
  COUNT(*) AS order_count
FROM `target_sql.orders` AS o
JOIN `target_sql.payments` AS p
ON  o.order_id = p.order_id
GROUP BY
  order_year,
  order_month,
  payment_type
ORDER BY
  order_year,
  order_month,
  payment_type;
```

Row	order_year	order_month	payment_type	order_count
1	2016	9	credit_card	3
2	2016	10	UPI	63
3	2016	10	credit_card	254
4	2016	10	debit_card	2
5	2016	10	voucher	23
6	2016	12	credit_card	1
7	2017	1	UPI	197
8	2017	1	credit_card	583
9	2017	1	debit_card	9

From this query we can get these following insights:

TARGET's Brazil E-com Market EDA on Google Big Query

Analysis By: Suman Kumar Nandi

1. Voucher Payments: in the data voucher payments remains relatively stable , with minor fluctuations in the count of orders.

2. Debit card payments: debit card payments showed a consistent count of orders, with no prominent variation or trend observe.

3. UPI Payments: UPI shows fluctuations and occasional peaks, there were notable peaks in UPI payments in October 2016, January 2017, May 2017, October 2017, Jan 2018, and May 2018.

4. Credit card payments: the count of an order made using credit cards shows fluctuations throughout the analyzed period, there was a gradual increase in credit card payments from Sep 2016 to Dec 2016.
The count experienced fluctuations in the following months, with a peck in May 2017, October 2017, January 2018, and May 2018

Recommendations
1. Capitalize on credit card peaks:
2. Harness the potential of upi payments
3. Evaluate Voucher Strategy
4. Monitor Debit card usage

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

```sql
SELECT
  payment_installments,
  COUNT(*) AS order_count
FROM
  `target-sql-423910.target_sql.orders`as o
JOIN `target_sql.payments` AS P
ON o.order_id = P.order_id
GROUP BY
  payment_installments
ORDER BY
  payment_installments;
```

TARGET’s Brazil E-com Market EDA on Google Big Query

Analysis By: Suman Kumar Nandi

Row	payment_installment	order_count
1	0	2
2	1	52546
3	2	12413
4	3	10461
5	4	7098
6	5	5239
7	6	3920
8	7	1626

Insights:

1. Most people tend to pay in single-payment installments.
2. 2-3 installments count in moderate
3. Mostly no orders are between 0 to 9

Recommendations:

1. Flexible payment installment option
2. Promote split payment option
3. Personalized payment recommendation
4. Transparent Pricing and Terms

By implementing this recommendation, businesses can enhance the payment experience, accommodate, various customers' preferences, and potentially increase the count of orders across
different payment installment options.