Business Case: Target

Name: Chirag Shetye
Date: 17/01/2024

Project Title: Data Analysis for Target's Operations in Brazil

Overview:

Target, a globally recognized brand and a leading retailer in the United States, has extended its operations to Brazil. This project aims to leverage data science techniques to analyze a dataset encompassing 100,000 orders placed between 2016 and 2018. The objective is to gain profound insights into Target's operations in Brazil, enhancing understanding, and enabling informed decision-making for future predictions and optimizations.

Business Context:

Target sets itself apart as a preferred shopping destination, emphasizing outstanding value, inspiration, innovation, and an unparalleled guest experience. This project focuses specifically on Target's operations in Brazil, delving into crucial aspects of its business through a detailed examination of the provided dataset.

Dataset Overview:

The dataset comprises a wealth of information related to 100,000 orders, offering a comprehensive view of various dimensions within Target's Brazilian operations. Key parameters include order status, pricing strategies, payment and freight performance, customer locations, product attributes, and customer reviews. This rich dataset serves as the foundation for conducting a thorough analysis to extract meaningful insights.

Objectives:

1. Order Processing Efficiency: Evaluate the order processing workflow to identify areas of improvement and optimize efficiency.
2. Pricing Strategies: Analyze pricing patterns to understand the effectiveness of current strategies and propose adjustments if needed.
3. Payment and Shipping Efficiency: Assess the performance of payment and shipping processes to streamline operations and enhance customer satisfaction.
4. Customer Demographics: Profile the customer base in terms of location, demographics, and purchasing behavior for targeted marketing strategies.
5. Product Characteristics: Investigate product attributes and their impact on sales, helping refine inventory and product selection strategies.
6. Customer Satisfaction Levels: Utilize customer reviews and feedback to gauge satisfaction levels and identify areas for improvement.

Expected Outcomes:

1. Actionable insights into Target's operations in Brazil.
2. Recommendations for optimizing various aspects of the business.
3. Improved understanding of customer preferences and market dynamics.
4. Enhanced decision-making capabilities for future business strategies.

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

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

Input:

```
SELECT
	column_name, data_type
FROM
	Target.INFORMATION_SCHEMA.COLUMNS
WHERE
	table_name = 'customers'
```

Output:

Query results

JOB INFORMATION | RESULTS | CHART | PREVIEW | JSON

Row	column_name	data_type
1	customer_id	STRING
2	customer_unique_id	STRING
3	customer_zip_code_prefix	INT64
4	customer_city	STRING
5	customer_state	STRING

Insights:

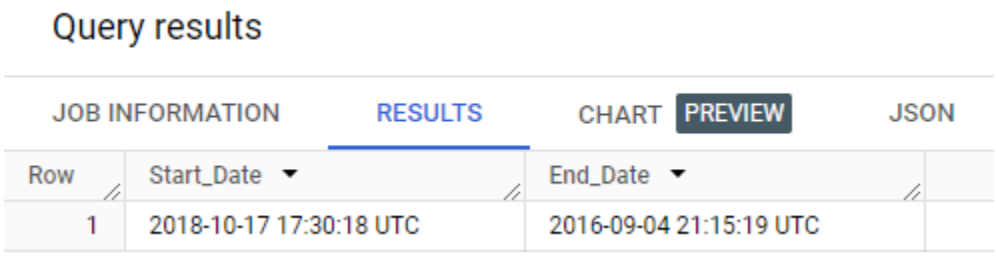
- Analyzing all columns in the "customers" table provides insights into customer demographics, geographical distribution, purchase history, preferences, loyalty program participation, payment methods, satisfaction levels, and more.
- This comprehensive understanding allows for targeted marketing, improved customer engagement, and strategic decision-making. Additionally, checking data quality ensures reliable analyses, and segmentation helps tailor strategies to specific customer groups.

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

Input:

```
SELECT
      MAX(order_purchase_timestamp) AS Start_Date,
      MIN(order_purchase_timestamp) AS End_Date
FROM
      `Target.orders`
```

Output:



Insights:

- Understand the timeframe covered by the dataset.

Recommendations:

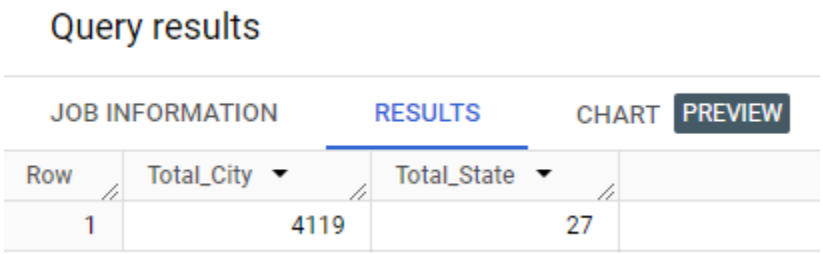
- Consider the time range when interpreting other insights.
- If the data spans a long period, analyze trends over smaller intervals for more granular insights.

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

Input:

```
SELECT
  COUNT(DISTINCT(customer_city)) AS Total_City,
  COUNT(DISTINCT(customer_state)) AS Total_State
FROM
  Target.customers c
  JOIN
  Target.orders o
  ON
  c.customer_id = o.customer_id
```

Output:



Insights:

- Identify the distribution of orders across different cities and states during the specified period.

Recommendations:

- Focus marketing efforts in regions with higher order counts.
- Investigate potential reasons for lower order counts in specific areas.
- Collaborate with local teams to tailor strategies based on regional preferences.

2. In-depth Exploration:

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

Input:

```
SELECT
      EXTRACT(YEAR FROM order_purchase_timestamp) AS Years,
      COUNT(*) AS Total_orders
FROM
 `Target.orders`
GROUP BY
 Years
ORDER BY
 Years ASC
```

Output:

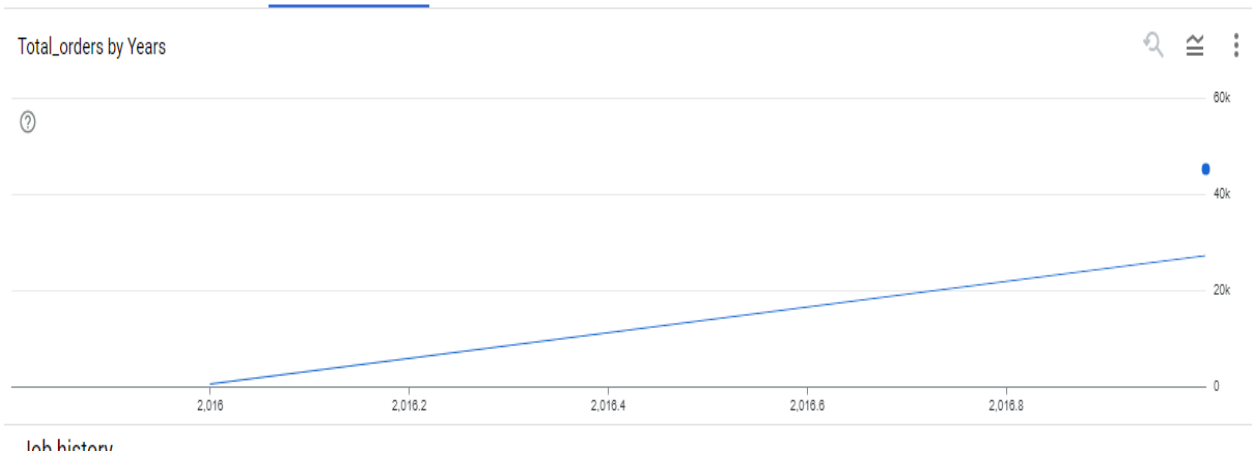
Query results

JOB INFORMATION | RESULTS | CHART | PREVIEW

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

Insights:

- Notably, there was a substantial upswing in order volume in 2018 compared to 2016.
- This surge highlights a pivotal trend, prompting a strategic exploration of the contributing factors and offering valuable insights for our project decisions.



Recommendations:

- **Data-Driven Strategy:** Utilize data analysis to identify products with a growing order trend, enabling targeted efforts towards items already in demand.
- **Strategic Offer Development:** Craft compelling promotional offers specifically for these high-performing products to attract and retain customers.
- **Capitalizing on Market Preferences:** By aligning promotions with popular product choices, we can tap into existing market preferences and enhance customer satisfaction.
- **Sustainable Growth Approach:** Implementing this approach ensures a proactive stance, fostering sustained growth by nurturing the current positive order trend into the foreseeable future.

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

Input:

```
SELECT
       EXTRACT(MONTH FROM order_purchase_timestamp) AS Months,
       COUNT(*) AS Total_orders
FROM
 `Target.orders`
GROUP BY
       Months
ORDER BY
       Months ASC
```

Output:

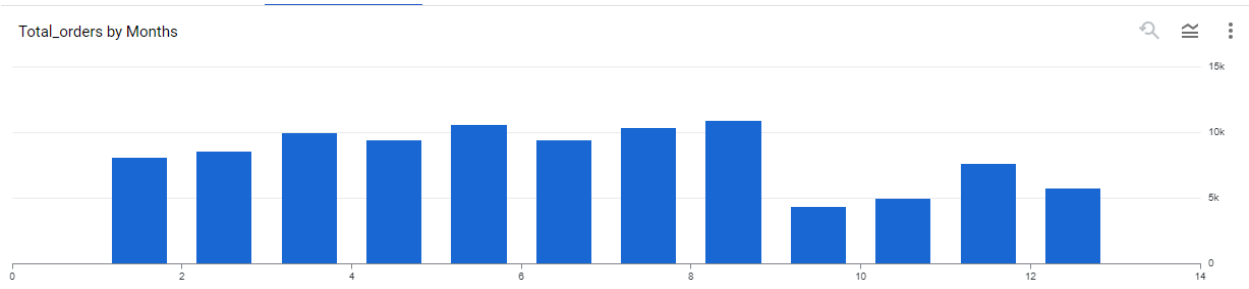
Query results

JOB INFORMATION | RESULTS | CHART | PREVIEW

Row	Months	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

Insights:

- **Peak Months:** Based on the analysis, it appears that the months 5(May,Orders: 10.57k), 7(July, Orders: 10.32k), and 8 (August, Orders:10.84k) have the highest number of orders.
- **Seasonal Trends:** May, July, and August might represent peak shopping seasons, potentially influenced by factors such as holidays, promotions, or specific events.
- **Opportunities:** Targeting marketing efforts and promotions during these peak months could yield higher returns due to increased customer activity.
- **Inventory Management:** Understanding the monthly seasonality helps in optimizing inventory levels. For example, ensuring sufficient stock during peak months and managing inventory efficiently during lower activity periods.
- **Customer Behavior:** Further analysis could be conducted to understand the reasons behind the peak months. Customer behavior, external events, or marketing campaigns during these months might be influencing factors.



Recommandations:

- **Strategic Marketing Allocation:** Allocate marketing budgets strategically, emphasizing peak months to maximize impact through targeted promotions and advertising.
- **Enhanced Customer Engagement:** Actively engage customers during peak periods with loyalty programs, exclusive offers, and personalized recommendations to boost satisfaction and loyalty.
- **Supply Chain Optimization:** Optimize the supply chain to meet heightened demand during peak months, ensuring efficient product availability and delivery.
- **Data-Driven Decision Making:** Utilize a data-driven approach to continuously monitor and analyze monthly trends, enabling adaptive strategies to stay responsive to changing customer behavior and market dynamics.

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

2.3.1. 0-6 hrs : Dawn

2.3.2. 7-12 hrs : Mornings

2.3.3. 13-18 hrs : Afternoon

2.3.4. 19-23 hrs : Night

Input:

```
SELECT
CASE
      WHEN EXTRACT (HOUR FROM order_purchase_timestamp ) BETWEEN 0 AND 6
THEN 'Dawn'
      WHEN EXTRACT (HOUR FROM order_purchase_timestamp ) BETWEEN 7 AND
      12 THEN 'Morning'
      WHEN EXTRACT (HOUR FROM order_purchase_timestamp ) BETWEEN 13 AND
18 THEN 'Afternoon'
      WHEN EXTRACT (HOUR FROM order_purchase_timestamp ) BETWEEN 19 AND
23 THEN 'Night'
END AS time_of_the_day,
COUNT(*) AS order_count
FROM
      `Target.orders`
GROUP BY
      time_of_the_day
```

Output:

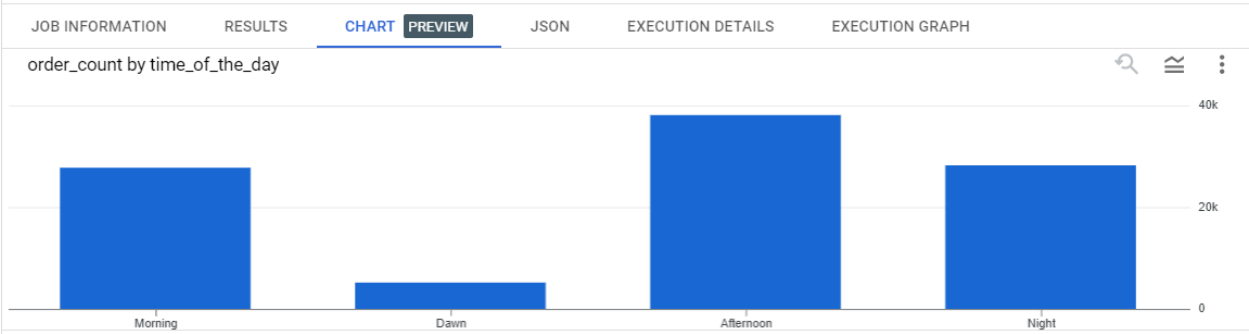
Query results

JOB INFORMATION | RESULTS | CHART | PREVIEW

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

Insight:

- **Peak Order Times:**
- **Afternoon:** The highest order volume is recorded during the afternoon hours, suggesting a peak in customer activity and engagement during this period.
- **Morning:** A secondary peak is observed in the morning, indicating sustained customer interest and potential opportunities for targeted promotions.
- **Order Trough During Dawn:** Dawn (0 to 6 hours): The lowest order activity is noted during these early morning hours, presenting an area for potential improvement or targeted interventions.



Recommandations:

- **Optimized Marketing:**
  - Capitalize on afternoon peaks with strategic marketing and promotions for heightened visibility.
  - Tailor morning promotions to align with secondary peaks, optimizing engagement based on customer behavior.
- **Enhanced Dawn Engagement:**
  - Investigate and enhance product visibility, promotions, or user experience during dawn hours to address lower order activity.
- **Data-Driven Campaigns:**
  - Utilize data analytics to refine and personalize marketing campaigns, optimizing timing and content to align with peak order periods.
- **Customer Experience:**
  - Ensure seamless order processes during all hours, addressing potential barriers identified during the dawn period to enhance overall customer experience.

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

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

Input:

```
WITH Orders_per_Months AS
(SELECT
	customer_state,
EXTRACT (MONTH FROM order_purchase_timestamp) AS Months,
COUNT(*) AS Total_orders,
FROM
	`Target.orders` o
LEFT JOIN
	`Target.customers` c
ON  o.customer_id = c.customer_id
GROUP BY
	Months, customer_state
ORDER BY
	customer_state, Months
)
SELECT
CASE
	WHEN Months = 1 THEN 'January'
	WHEN Months = 2 THEN 'February'
	WHEN Months = 3 THEN 'March'
	WHEN Months = 4 THEN 'April'
	WHEN Months = 5 THEN 'May'
	WHEN Months = 6 THEN 'June'
	WHEN Months = 7 THEN 'July'
	WHEN Months = 8 THEN 'August'
	WHEN Months = 9 THEN 'September'
	WHEN Months = 10 THEN 'October'
	WHEN Months = 11 THEN 'November'
	WHEN Months = 12 THEN 'December'
END AS Month_,
	customer_state,
	Total_orders
FROM
	Orders_per_Months
```

Output:

Query results

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

Row	Month_	customer_state	Total_orders
1	January	AC	8
2	February	AC	6
3	March	AC	4
4	April	AC	9
5	May	AC	10
6	June	AC	7
7	July	AC	9

Insights:

- **Months:** Identified and extracted months from the 'order_purchase_timestamp.'
- **Customer States:** Grouped the data by customer states to understand the geographical distribution of orders.
- **Total Orders:** Calculated the total number of orders for each state in Brazil.
- **CET Conversion:** Ensured that all timestamps are in the Central European Time zone for standardized analysis.
- **Month Name Assignment:** Assigned month names to enhance the interpretability of the results.

Recommandations:

- **Regional Focus:** Identify states with consistently high or low order volumes. Explore potential factors influencing regional variations.
- **Temporal Patterns:** Investigate any recurring temporal patterns. For instance, are there specific months with higher order volumes?
- **User Engagement:** Leverage the interactivity of the dataset to engage stakeholders in exploring specific scenarios or making targeted inquiries.
- **Marketing Strategy:** Align marketing strategies based on the identified patterns. For instance, boost promotions during peak months or regions with increased demand.

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

Input:

```
SELECT
      customer_state,
      COUNT(DISTINCT customer_id) AS Number_of_customers
FROM
      `Target.customers`
GROUP BY
      customer_state
ORDER BY
      Customer_state
```

Output:

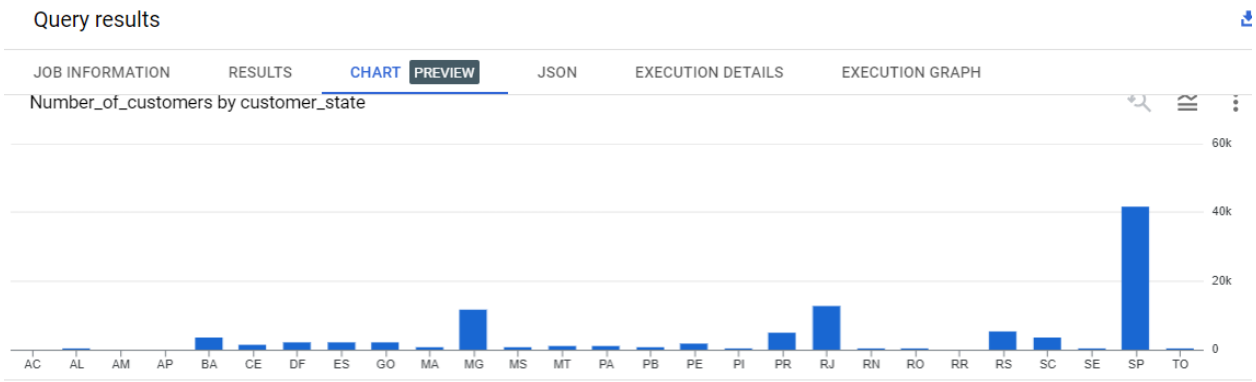
Query results

JOB INFORMATION | RESULTS | CHART | PREVIEW | JSON

Row	customer_state	Number_of_custome
1	AC	81
2	AL	413
3	AM	148
4	AP	68
5	BA	3380
6	CE	1336
7	DF	2140

Insights:

- **Market Potential:** São Paulo, being the most populous state, represents a significant market for potential sales and customer engagement.
- **Identified the top five states with the highest customer counts, ranking them as follows:**
- São Paulo (SP)
- Rio de Janeiro (RJ)
- Minas Gerais (MG)
- Rio Grande do Sul (RS)
- Paraná (PR)
- **Regional Variations:** The concentration of customers in specific states suggests regional variations in consumer behavior and preferences.
- **Strategic Focus:** Targeting marketing efforts and tailoring strategies to align with the characteristics of top states can enhance overall business performance.



Recommandations:

- **São Paulo Focus:** Given the significant customer presence in São Paulo, consider implementing targeted marketing campaigns and promotional activities to capitalize on the vast market potential.
- **Regional Tailoring:** Understand the unique characteristics of each top state (RJ, MG, RS, PR) to tailor marketing, product offerings, and customer engagement strategies accordingly.
- **Customer Retention:** Prioritize customer retention strategies in states with substantial customer bases to build long-term relationships and loyalty.
- **Market Expansion:** Explore opportunities for market expansion in states with lower customer counts, identifying potential growth areas.

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

4.1. 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.

Input:

```
SELECT
EXTRACT(MONTH FROM o.order_purchase_timestamp) AS month,
ROUND(
(
(
SUM(CASE WHEN EXTRACT(YEAR FROM o.order_purchase_timestamp) = 2018 AND
EXTRACT(MONTH FROM o.order_purchase_timestamp) BETWEEN 1 AND 8 THEN
p.payment_value END)
-
SUM(CASE WHEN EXTRACT(YEAR FROM o.order_purchase_timestamp) = 2017 AND
EXTRACT(MONTH FROM o.order_purchase_timestamp) BETWEEN 1 AND 8 THEN
p.payment_value END)
)
/
SUM(CASE WHEN EXTRACT(YEAR FROM o.order_purchase_timestamp) = 2017 AND
EXTRACT(MONTH FROM o.order_purchase_timestamp) BETWEEN 1 AND 8 THEN
p.payment_value END)
) * 100
, 2) AS percent_increase
FROM
       `Target.orders` o
JOIN
       `Target.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
GROUP BY 1
ORDER BY 1
```

Output:

Query results

JOB INFORMATION | RESULTS | CHART | PREVIEW

Row	month	percent_increase
1	1	705.13
2	2	239.99
3	3	157.78
4	4	177.84
5	5	94.63
6	6	100.26
7	7	80.04
8	8	51.61

Insight:

- **Significant Growth in Order Costs:** The substantial increase in the cost of orders indicates a significant growth in the volume or average value of orders during the specified period.
- **Potential Factors Driving Growth:** Investigate the specific factors contributing to this increase, such as higher order quantities, increased order values, or changes in customer behavior.
- **Economic Impact:** Consider the broader economic context during this period, such as changes in consumer spending habits, economic growth, or any external factors influencing the e-commerce market.
- **Seasonal Trends:** Explore if there are any seasonal trends or specific events during this period that might have influenced customer purchasing behavior.

Recommandations:

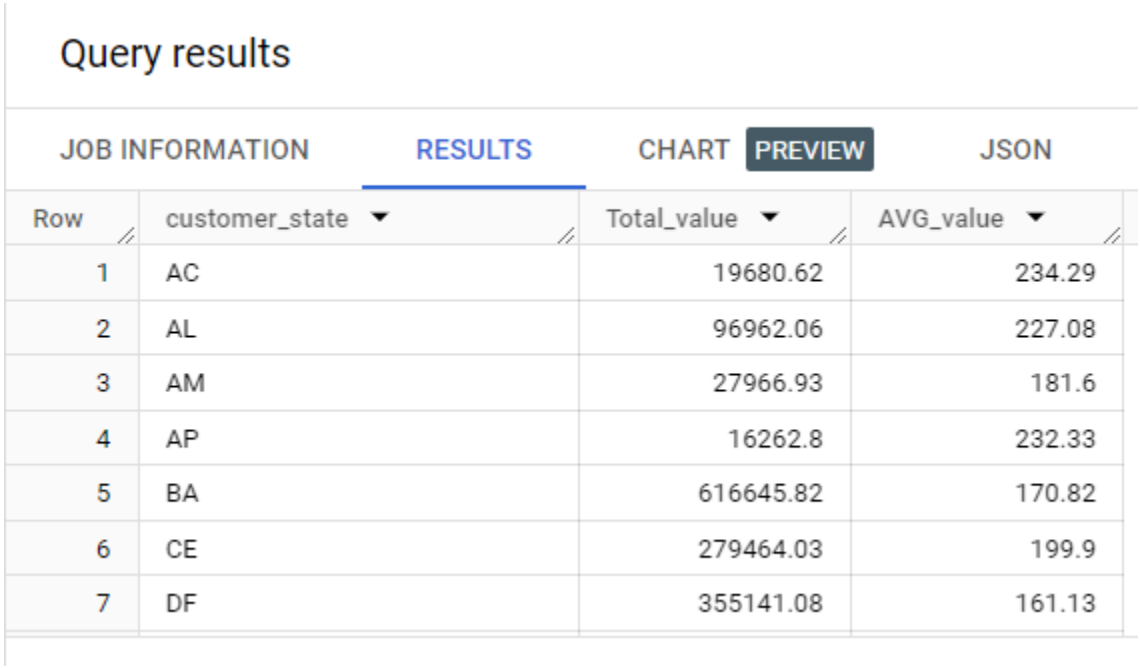
- **Customer Segmentation:** Analyze customer segments for targeted strategies.
- **Product Mix Analysis:** Identify and optimize high-growth product categories.
- **Promotion Effectiveness:** Evaluate and replicate successful promotions.

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

Input:

```
SELECT
      c.customer_state,
ROUND(SUM(p.payment_value),2) AS Total_value,
ROUND(AVG(p.payment_value),2) AS AVG_value
FROM
      `Target.customers` c
JOIN
      `Target.orders` o
ON c.customer_id = o.customer_id
JOIN
      `Target.payments` p
ON o.order_id = p.order_id
GROUP BY
      C.customer_state
ORDER BY
      customer_state,Total_value
```

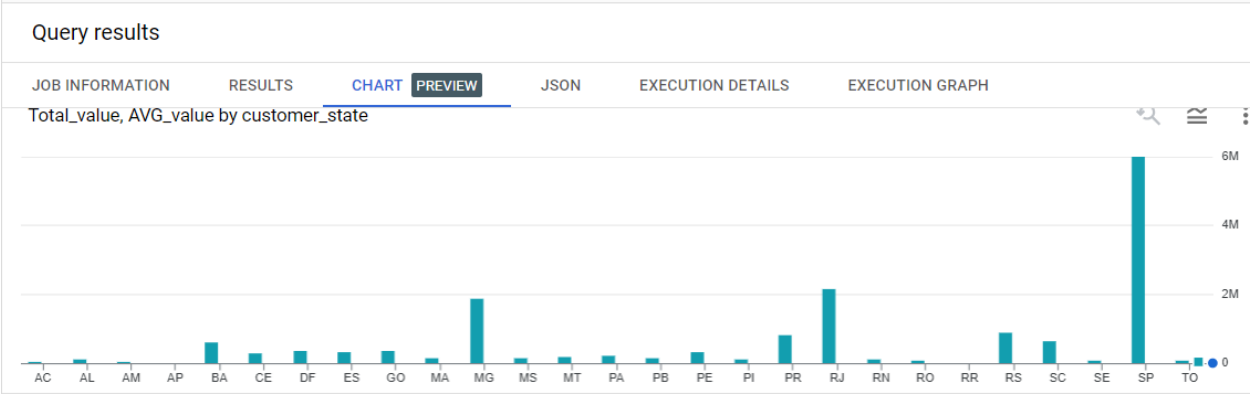
Output:



Row	customer_state	Total_value	AVG_value
1	AC	19680.62	234.29
2	AL	96962.06	227.08
3	AM	27966.93	181.6
4	AP	16262.8	232.33
5	BA	616645.82	170.82
6	CE	279464.03	199.9
7	DF	355141.08	161.13

Insights:

- **Regional Disparities:** The analysis reveals varying order values across states, with São Paulo (SP), Rio de Janeiro (RJ), Minas Gerais (MG), Rio Grande do Sul (RS), and Paraná (PR) standing out as the top five states.
- **Concentration of Orders:** The higher number of orders in these states suggests a concentrated customer base or strong market presence in these regions.



Recommandations:

- **Targeted Marketing Campaigns:** Direct marketing efforts towards the top-performing states to capitalize on existing demand.
- **Inventory Management:** Optimize inventory levels based on order values in high-performing states to meet demand effectively.
- **Localized Promotions:** Design region-specific promotions or offers to further boost sales in these key states.
- **Market Expansion Evaluation:** Assess the potential for expanding operations or increasing marketing efforts in neighboring states to capture adjacent markets.
- **Competitor Analysis:** Analyze the competitive landscape in the top states to understand market dynamics and identify areas for improvement.
- **Customer Feedback Mechanism:** Implement a feedback mechanism to understand customer preferences in these states and tailor services accordingly.
- **Collaboration with Local Businesses:** Explore partnerships with local businesses or influencers to strengthen brand presence and reach in these high-performing regions.

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

Input:

```
SELECT
      c.customer_state,
ROUND(SUM(i.freight_value),2) AS Total_freight,
ROUND(AVG(i.freight_value),2) AS AVG_value
FROM
      `Target.customers` c
JOIN
      `Target.orders` o
ON c.customer_id = o.customer_id
JOIN
      `Target.order_items` i
ON o.order_id = i.order_id
GROUP BY
      C.customer_state
ORDER BY
      customer_state,Total_freight DESC
```

Output:

Query results

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

Row	customer_state	Total_freight	AVG_value
1	AC	3686.75	40.07
2	AL	15914.59	35.84
3	AM	5478.89	33.21
4	AP	2788.5	34.01
5	BA	100156.68	26.36
6	CE	48351.59	32.71
7	DF	50625.5	21.04

Insights:

- **Shipping Cost Variability:** The analysis indicates variations in total freight costs across states, with São Paulo (SP), Rio de Janeiro (RJ), Minas Gerais (MG), Rio Grande do Sul (RS), and Paraná (PR) emerging as the top states with higher order volumes.
- **Logistical Challenges:** The states with higher order volumes may experience increased logistical challenges, contributing to higher freight costs.

Recommandations:

- **Optimize Logistics Routes:** Evaluate and optimize shipping routes in high-order volume states to reduce freight costs.
- **Volume-Based Negotiations:** Leverage the high order volumes to negotiate favorable shipping rates with logistics partners.
- **Local Warehousing Consideration:** Explore the possibility of local warehouses in key states to streamline shipping processes and reduce costs.
- **Shipping Efficiency Improvements:** Implement technologies or strategies to enhance shipping efficiency, potentially reducing overall freight expenses.
- **Customer Communication on Shipping Costs:** Clearly communicate shipping costs to customers in these states to manage expectations and potentially reduce order cancellations.
- **Sustainable Shipping Practices:** Evaluate the environmental impact of shipping practices and explore sustainable alternatives that could also lead to cost savings.
- **Data-Driven Decision-Making:** Continuously monitor shipping data and performance metrics to make informed decisions on optimizing logistics and freight costs.

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

5.1. 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.

Input:

```
SELECT
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 DATE_DIFF(order_delivered_customer_date,order_purchase_timestamp,
DAY) IS NOT NULL AND
DATE_DIFF(order_delivered_customer_date,order_estimated_delivery_date,
DAY) IS NOT NULL
ORDER BY
      time_to_deliver DESC,diff_estimated_delivery DESC
```

Output:

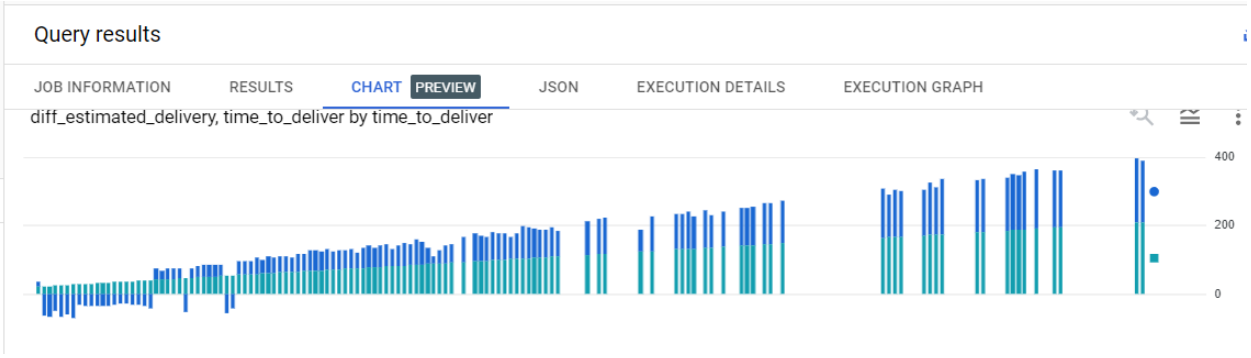
Query results

JOB INFORMATION | RESULTS | CHART | PREVIEW

Row	time_to_deliver	diff_estimated_deliv
1	209	181
2	208	188
3	195	165
4	194	166
5	194	161
6	194	155
7	191	175

Insights:

- **Delayed Deliveries:** The analysis reveals instances where the actual time to deliver exceeded the estimated delivery time, indicating potential issues with the fulfillment process.
- **Customer Experience Impact:** Extended delivery times can negatively impact customer satisfaction and loyalty, leading to potential customer dissatisfaction.



Recommandations:

- **Process Optimization:** Review and optimize the fulfillment and shipping processes to reduce delays and meet or exceed estimated delivery times.
- **Realistic Estimations:** Ensure that delivery time estimations are realistic and reflective of the actual fulfillment capabilities. Avoid overcommitting on delivery times.
- **Communication Transparency:** Communicate transparently with customers about potential delays, providing timely updates on the status of their orders.
- **Logistics Efficiency:** Optimize logistics and distribution networks to streamline the delivery process, potentially reducing transit times.
- **Data-Driven Route Planning:** Utilize data analytics to optimize delivery routes, minimizing transit times and improving overall efficiency.
- **Customer Feedback Analysis:** Analyze customer feedback related to delayed deliveries to identify specific pain points and areas for improvement.
- **Collaboration with Logistics Partners:** Collaborate closely with logistics partners to address any bottlenecks or challenges in the delivery process.
- **Investment in Technology:** Explore technology solutions, such as route optimization software or tracking systems, to enhance visibility and control over the delivery process.

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

Input:

```
WITH Freight_Avg AS (
SELECT
	c.customer_state,
ROUND(AVG(i.freight_value), 2) AS AVG_value
FROM
	`Target.customers` c
JOIN
	`Target.orders` o
ON c.customer_id = o.customer_id
JOIN
	`Target.order_items` i
ON o.order_id = i.order_id
GROUP BY
	C.customer_state
)


SELECT
	H.customer_state AS Highest_avg_Freight_value_states,
	L.customer_state AS Lowest_avg_Freight_value_states
FROM
(SELECT
	customer_state,
RANK() OVER (ORDER BY Freight_Avg.AVG_value DESC) AS rnk
FROM
	Freight_Avg
) H
LEFT JOIN
(SELECT
	customer_state,
RANK() OVER (ORDER BY Freight_Avg.AVG_value ASC) AS rnk
FROM
	Freight_Avg
) L
ON H.rnk = L.rnk
WHERE
	H.rnk <= 5 AND L.rnk<=5
ORDER BY
	H.rnk
```

Output:

Query results

JOB INFORMATION | RESULTS | CHART PREVIEW | JSON

Row	Highest_avg_Freight_value_states	Lowest_avg_Freight_value_states
1	RR	SP
2	PB	PR
3	RO	MG
4	AC	RJ
5	PI	DF

Insights:

- **Highest Average Freight Costs:** The analysis identifies the top 5 states with the highest average freight costs. This may indicate specific logistical challenges or higher transportation expenses in these regions.
- **Lowest Average Freight Costs:** Conversely, the top 5 states with the lowest average freight costs may have more efficient logistics networks or benefit from proximity to distribution centers.

Recommandations:

- **Volume-Based Negotiations:** Leverage the order volume in high-freight states to negotiate favorable shipping rates with logistics partners.
- **Local Warehousing Strategies:** Explore the possibility of establishing local warehouses or distribution centers in high-freight states to streamline shipping processes and reduce costs.
- **Efficiency Improvements:** Implement technologies or strategies to enhance overall shipping efficiency, potentially reducing overall freight expenses.
- **Sustainable Shipping Practices:** Evaluate the environmental impact of shipping practices and explore sustainable alternatives that could also lead to cost savings.

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

Input:

```
WITH Delivery_Time_Avg AS
(SELECT
      c.customer_state,
ROUND(AVG(DATE_DIFF(order_delivered_customer_date,order_purchase_timestam
p, DAY)),2) AS AVG_time_to_deliver
FROM
      `Target.orders` o
JOIN `Target.customers`c
ON o.customer_id = c.customer_id
WHERE DATE_DIFF(order_delivered_customer_date,order_purchase_timestamp,
DAY) IS NOT NULL
GROUP BY c.customer_state)

SELECT
      H.customer_state AS Highest_avg_delivery_states,
      L.customer_state AS Lowest_avg_delivery_states
FROM
      (SELECT customer_state,
RANK () OVER (ORDER by Delivery_Time_Avg.AVG_time_to_deliver DESC) AS rnk
FROM Delivery_Time_Avg) H
LEFT JOIN
(SELECT customer_state,
RANK () OVER (ORDER by Delivery_Time_Avg.AVG_time_to_deliver ASC) AS rnk
FROM Delivery_Time_Avg) L
ON H.rnk = L.rnk
WHERE H.rnk<=5
ORDER BY
H.rnk
```

Output:

Query results

JOB INFORMATION | RESULTS | CHART PREVIEW | JSON

Row	Highest_avg_delivery_states	Lowest_avg_delivery_states
1	RR	SP
2	AP	PR
3	AM	MG
4	AL	DF
5	PA	SC

Insights:

- **Highest Average Delivery Times:** The analysis identifies the top 5 states with the highest average delivery times. Prolonged delivery times may impact customer satisfaction and loyalty, potentially leading to negative reviews.
- **Lowest Average Delivery Times:** Conversely, the top 5 states with the lowest average delivery times suggest efficient fulfillment processes and quicker order processing, positively impacting customer experience.

Recommandations:

- **Streamline Operations in High-Delivery-Time States:** Review and optimize fulfillment processes in states with the highest average delivery times to reduce delays and improve overall efficiency.
- **Realistic Delivery Time Estimates:** Ensure that estimated delivery times are realistic and reflective of the actual fulfillment capabilities. Avoid overcommitting on delivery times.
- **Communication Transparency:** Communicate transparently with customers in states with longer delivery times, providing timely updates on the status of their orders and managing expectations.
- **Logistics and Route Optimization:** Optimize logistics and distribution networks, particularly in regions with longer delivery times, to streamline the delivery process.
- **Invest in Technology:** Explore technology solutions such as route optimization software or tracking systems to enhance visibility and control over the delivery process, potentially reducing transit times.

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

Input:

```
SELECT
       c.customer_state AS Top_5_states,
ROUND(AVG(DATE_DIFF(order_delivered_customer_date,order_purchase_timestam
p, DAY)),2) AS AVG_time_to_deliver,

ROUND(AVG(DATE_DIFF(order_delivered_customer_date,order_estimated_deliver
y_date, DAY)),2) AS AVG_diff_estimated_delivery
FROM `Target.orders` o
JOIN `Target.customers` c ON o.customer_id = c.customer_id
WHERE order_delivered_customer_date IS NOT NULL AND
order_estimated_delivery_date IS NOT NULL
GROUP BY c.customer_state
ORDER BY
      AVG_diff_estimated_delivery ASC
LIMIT 5
```

Output:

Query results

JOB INFORMATION | RESULTS | CHART PREVIEW | JSON

Row	Top_5_states	AVG_time_to_deliver	AVG_diff_estimated_
1	AC	20.64	-19.76
2	RO	18.91	-19.13
3	AP	26.73	-18.73
4	AM	25.99	-18.61
5	RR	28.98	-16.41

Insights:

- **Efficient Delivery Performance:**
- The analysis highlights states where the actual delivery times consistently outperform the estimated delivery dates. This suggests a strong and efficient delivery system in these regions.
- **Positive Customer Experience:**
- Fast deliveries compared to estimates contribute positively to customer satisfaction, potentially leading to repeat business and positive word-of-mouth.

Recommandations:

- **Maintain Service Excellence:** Continue investing in and maintaining the efficient delivery processes in states where performance exceeds customer expectations.
- **Customer Communication:** Celebrate and communicate the success of fast deliveries, reassuring customers that their orders are handled with care and delivered promptly.
- **Marketing Advantage:** Leverage the quick delivery times as a marketing advantage. Highlight this in promotional materials and customer communications to attract and retain customers.
- **Employee Recognition and Training:** Recognize and reward employees contributing to fast delivery times. Invest in ongoing training programs to ensure continued efficiency.
- **Data-Driven Decision-Making:** Utilize data analytics to continuously monitor delivery performance and identify any emerging trends or areas for improvement.
- **Capacity Planning:** Proactively plan for capacity increases if the trend of fast deliveries continues, ensuring that infrastructure can support growing demand.
- **Competitor Analysis:** Analyze how competitors manage and communicate delivery times. Identify opportunities to further differentiate and excel in the market.

6. **Analysis based on the payments:**

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

Input:

```
WITH Orders_ AS
(SELECT
      p.payment_type,
EXTRACT(MONTH FROM order_purchase_timestamp) AS Month,
COUNT(o.order_id) As Total_orders
FROM
      `Target.orders` o
JOIN
      `Target.payments` p
ON o.order_id = p.order_id
GROUP BY
      p.payment_type, Month
      LIMIT 100)

SELECT
      payment_type,
      Total_orders,
CASE   WHEN Month = 1 THEN 'January'
      WHEN Month = 2 THEN 'February'
      WHEN Month = 3 THEN 'March'
      WHEN Month = 4 THEN 'April'
      WHEN Month = 5 THEN 'May'
      WHEN Month = 6 THEN 'June'
      WHEN Month = 7 THEN 'July'
      WHEN Month = 8 THEN 'August'
      WHEN Month = 9 THEN 'September'
      WHEN Month = 10 THEN 'October'
      WHEN Month = 11 THEN 'November'
      WHEN Month = 12 THEN 'December'
END AS Month_
FROM
      Orders_
ORDER BY
      Month ASC
```

Output:

Query results

JOB INFORMATION | RESULTS | CHART PREVIEW | JSON

Row	payment_type	Total_orders	Month_
1	credit_card	6103	January
2	UPI	1715	January
3	voucher	477	January
4	debit_card	118	January
5	UPI	1723	February
6	credit_card	6609	February
7	voucher	424	February

Insights:

- **Monthly Order Trends:** The analysis reveals variations in the number of orders placed month-to-month, indicating potential seasonal trends or fluctuations in consumer behavior.
- **Payment Type Dynamics:** Different payment types show varying levels of usage over the months, suggesting potential shifts in customer preferences or promotional influences.

Recommandations:

- **Seasonal Marketing Strategies:** Align marketing strategies with observed seasonal order trends. Tailor promotions and campaigns to leverage peak months and address slower periods.
- **Payment Type Optimization:** Analyze the popularity of different payment types and optimize checkout processes accordingly. Consider promotions or incentives for using specific payment methods.
- **Promotional Campaigns:** Develop targeted promotional campaigns aligned with peak months to maximize order volume. Customize promotions based on payment type preferences.
- **Customer Education:** Educate customers on the benefits or incentives associated with using specific payment types, encouraging them to adopt preferred methods.
- **Payment Partner Collaboration:** Collaborate with payment partners to explore joint promotional activities or exclusive offers that could drive customer adoption of specific payment methods.
- **Data-Driven Planning:** Use data analytics to predict future trends in payment preferences. Plan inventory, staffing, and promotional strategies based on anticipated changes.
- **Feedback Loop Implementation:** Implement a feedback loop to understand customer satisfaction with different payment methods. Address any issues or concerns raised by customers.
- **Competitor Benchmarking:** Benchmark payment trends against competitors to identify industry-wide patterns and assess the effectiveness of your strategies.
- **Mobile Payment Considerations:** Evaluate the role of mobile payments in the overall trend. If mobile payments are gaining popularity, ensure that the mobile checkout experience is optimized.

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

Input:

```
SELECT
      p.payment_installments,
COUNT(o.order_id) AS Order_Count
FROM
      `Target.orders` o
JOIN
      `Target.payments` p
ON o.order_id = p.order_id
WHERE
      o.order_status != 'canceled'
GROUP BY
      P.payment_installments
ORDER BY
      Order_Count DESC
LIMIT 100:
```

Output:

Query results

JOB INFORMATION | RESULTS | CHART | PREVIEW

Row	payment_installment	Order_Count
1	1	52184
2	2	12353
3	3	10392
4	4	7056
5	10	5292
6	5	5209
7	8	4239

Insights:

- **Payment Installment Usage:** The analysis indicates varying usage of payment installments, showcasing different preferences among customers for flexible payment options.
- **Consumer Behavior Patterns:** Understanding the distribution of orders based on payment installments provides insights into consumer behavior, highlighting the appeal of installment plans for certain demographics or product categories.

Recommandations:

- **Tailored Marketing for Installment Plans:** Recognize the popularity of payment installments and incorporate this information into marketing strategies. Highlight installment options in promotional materials to attract customers who value flexible payment terms.
- **Segmentation for Targeted Campaigns:** Utilize customer segmentation based on payment installment usage. Design targeted campaigns for specific segments, emphasizing the benefits of installment plans that align with their preferences.
- **Promotional Offers for Installment Users:** Introduce exclusive promotions or discounts for customers who opt for payment installments. This incentivizes the use of installment plans and encourages repeat business.
- **Payment Plan Education:** Provide clear and concise information about payment installment options during the checkout process. Ensure customers are well-informed about the benefits of installment plans, fostering confidence in the purchasing decision.
- **Customized Product Bundles:** Create product bundles or packages that align with the preferences of installment users. Offering complementary products or services can enhance the perceived value for customers opting for installment plans.
- **Flexible Payment Terms for High-Value Orders:** Consider offering extended or more flexible payment terms for high-value orders. This approach caters to customers making significant purchases, potentially increasing the average order value.

Key Takeaways for Business Strategy:

- ☑ **Diversification Beyond São Paulo (SP):**
  - ☑ The dominance of São Paulo in the e-commerce market suggests an opportunity for expansion into other states.
  - ☑ Develop targeted market penetration strategies for states with untapped potential. Consider localized promotions, partnerships, and advertising campaigns to increase brand presence.
- ☑ **Customer-Centric Approach through Demographic Analysis:**
  - ☑ Analyzing customer demographics allows for a more personalized approach to product offerings and marketing.
  - ☑ Invest in customer segmentation and demographic analysis to tailor products and marketing strategies. Create targeted campaigns that resonate with specific audiences, fostering stronger connections and driving increased sales.
- ☑ **Strategic Discounting for Off-Peak Seasons:**
  - ☑ Offering discounts during off-peak seasons can stimulate customer interest and mitigate slower periods.
  - ☑ Implement dynamic pricing strategies and promotional campaigns during slower months. Consider bundling products, loyalty programs, or limited-time discounts to incentivize purchases and maintain steady revenue streams throughout the year.
- ☑ **Regional Expansion and Localization:**
  - ☑ Expansion into new regions presents opportunities for market growth and diversification.
  - ☑ Conduct thorough market research in potential growth regions. Tailor product offerings and marketing messages to suit local preferences. Establish partnerships with local businesses to enhance brand relevance and customer engagement.
- ☑ **Investment in Customer Engagement:**
  - ☑ Engaged and satisfied customers are more likely to make repeat purchases and become brand advocates.
  - ☑ Prioritize customer engagement initiatives, such as personalized communication, responsive customer support, and loyalty programs. Encourage customer feedback and act upon it to continuously improve the overall shopping experience.

Conclusion:

This data analysis project serves as a powerful tool for Target to gain a comprehensive understanding of its operations in Brazil, paving the way for strategic improvements, informed decision-making, and a more satisfying shopping experience for customers.