

Targeting Success:

A Case Study of Expansion into Brazil

-By Samarth Kolge

Overview of Target:

Target Corporation stands as one of the premier retail giants in the United States. Originating in 1902 as the Dayton Company, it rebranded to Target Corporation in 1962. Target distinguishes itself by offering stylish and affordable products, often collaborating with designers and brands to provide exclusive collections. With a robust online presence via its website and app, Target facilitates seamless online shopping, delivery and in-store pickup services, catering to the evolving needs of modern consumers.

Purpose of Case Study:

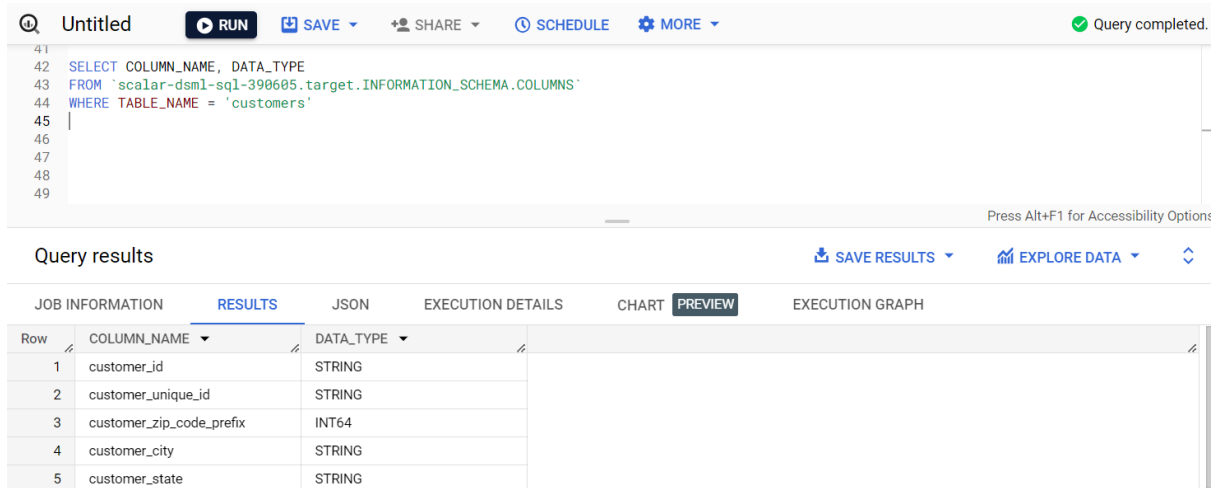
In its expansion endeavours, Target has ventured into the vibrant market of Brazil, sharing invaluable insights from a dataset comprising 100,000 orders spanning 2016 to 2018. The objective of this case study is to extract meaningful insights and provide actionable feedback based on the shared dataset. Through meticulous analysis and interpretation, we aim to unveil opportunities for optimization and strategic growth, driving Target's continued success in the dynamic landscape of retail.



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

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

Ans- `SELECT COLUMN_NAME, DATA_TYPE
FROM `scalar-dsml-sql-390605.target.INFORMATION_SCHEMA.COLUMNS`
WHERE TABLE_NAME = 'customers'`



The screenshot shows a SQL query editor interface. At the top, there's a toolbar with buttons for 'RUN', 'SAVE', 'SHARE', 'SCHEDULE', and 'MORE'. A status bar on the right indicates 'Query completed.' Below the toolbar, the SQL query is entered in a text area. The query is: `SELECT COLUMN_NAME, DATA_TYPE
FROM `scalar-dsml-sql-390605.target.INFORMATION_SCHEMA.COLUMNS`
WHERE TABLE_NAME = 'customers'`. Below the query editor, there's a section titled 'Query results' with tabs for 'JOB INFORMATION', 'RESULTS', 'JSON', 'EXECUTION DETAILS', 'CHART', 'PREVIEW', and 'EXECUTION GRAPH'. The 'RESULTS' tab is selected, showing a table with 5 rows and 2 columns: 'COLUMN_NAME' and 'DATA_TYPE'. The rows contain the following data:

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:**

Understanding datatypes is essential for correct interpretation and analysis. It helps ensure that we are working with appropriate manner.

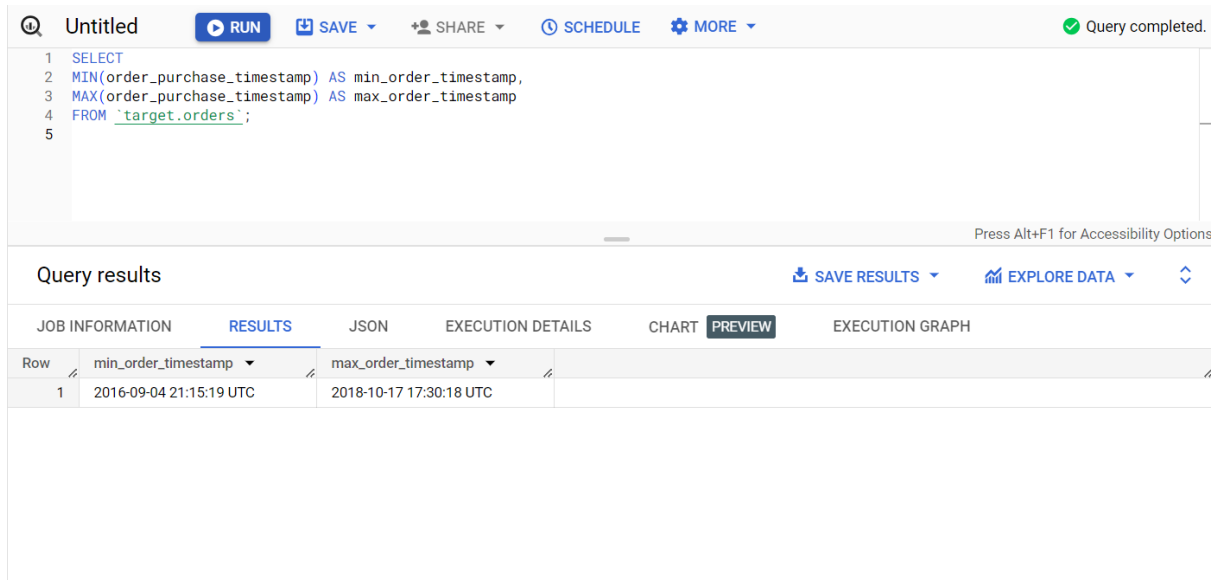
- **RECOMMENDATIONS:**

Ensure that datatype of the customer table matches nature of the data they contains. Incorrect datatypes may occur errors or inaccurate analysis.

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

Ans- `SELECT`

```
MIN(order_purchase_timestamp) AS min_order_timestamp,  
MAX(order_purchase_timestamp) AS max_order_timestamp  
FROM `target.orders`;
```



Query results

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

- **INSIGHTS :**

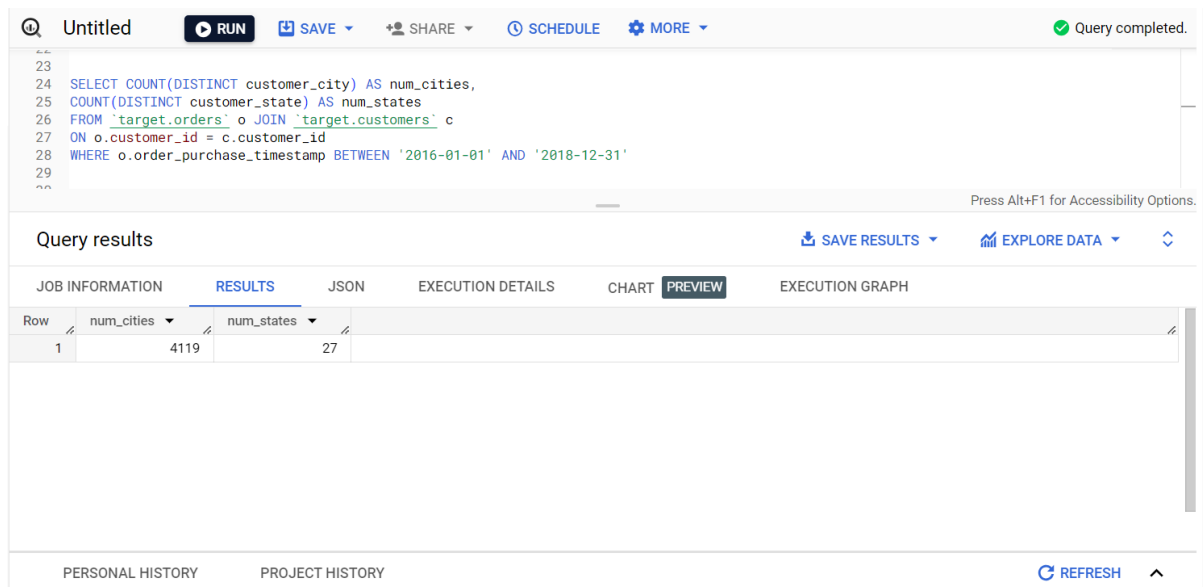
Target's order history spans a period of two years, beginning with their first order on **September 4th, 2016**, and concluding with their most recent order on **October 17th, 2018**. This indicates a consistent customer relationship over a significant duration.

- **RECOMMENDATIONS : N/A**

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

Ans-

```
SELECT COUNT(DISTINCT customer_city) AS num_cities,  
COUNT(DISTINCT customer_state) AS num_states  
FROM `target.orders` o JOIN `target.customers` c  
ON o.customer_id = c.customer_id  
WHERE o.order_purchase_timestamp BETWEEN '2016-01-01' AND '2018-12-31';
```



The screenshot shows a SQL query execution interface. The query is displayed in a text area, and the results are shown in a table below. The table has two columns: 'num_cities' and 'num_states'. The results show 4119 cities and 27 states.

Row	num_cities	num_states
1	4119	27

- **INSIGHTS :**

The dataset reveals that customers have placed orders from **4,199 cities across 27 states**. This demonstrates a widespread customer base and potential for further growth.

- **RECOMMENDATIONS :**

- To expand into new states and cities, we can introduce **exclusive offers** and implement targeted strategies to **attract new customers**.
- Conduct **advertisement campaigns** in cities where brand awareness is low. This will help capture the attention of potential customers and increase market reach.

2. In-depth Exploration:

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

Ans-

```
SELECT
EXTRACT(YEAR FROM order_purchase_timestamp) AS year,
COUNT(order_id) AS total_orders
FROM `target.orders`
GROUP BY year
ORDER BY year;
```

The screenshot shows a SQL query editor with the following query:

```
4 SELECT
5 EXTRACT(YEAR FROM order_purchase_timestamp) AS year,
6 COUNT(order_id) AS total_orders
7 FROM `target.orders`
8 GROUP BY year
9 ORDER BY year;
```

Below the query editor, the query results are displayed in a table. The table has two columns: 'year' and 'total_orders'. The results show a steady increase in orders over the years 2016, 2017, and 2018.

Row	year	total_orders
1	2016	329
2	2017	45101
3	2018	54011

- **INSIGHTS -**

- The order trend shows a **steady increase** over the past three years:
 - i. In **2016**, customers placed **329 orders**.
 - ii. In **2017**, the number jumped significantly to **45,101 orders**.
 - iii. In **2018**, the trend continued upward with **54,011 orders**.
 - iv. This indicates that the company has implemented effective strategies to drive growth, leading to a significant rise in order volume.

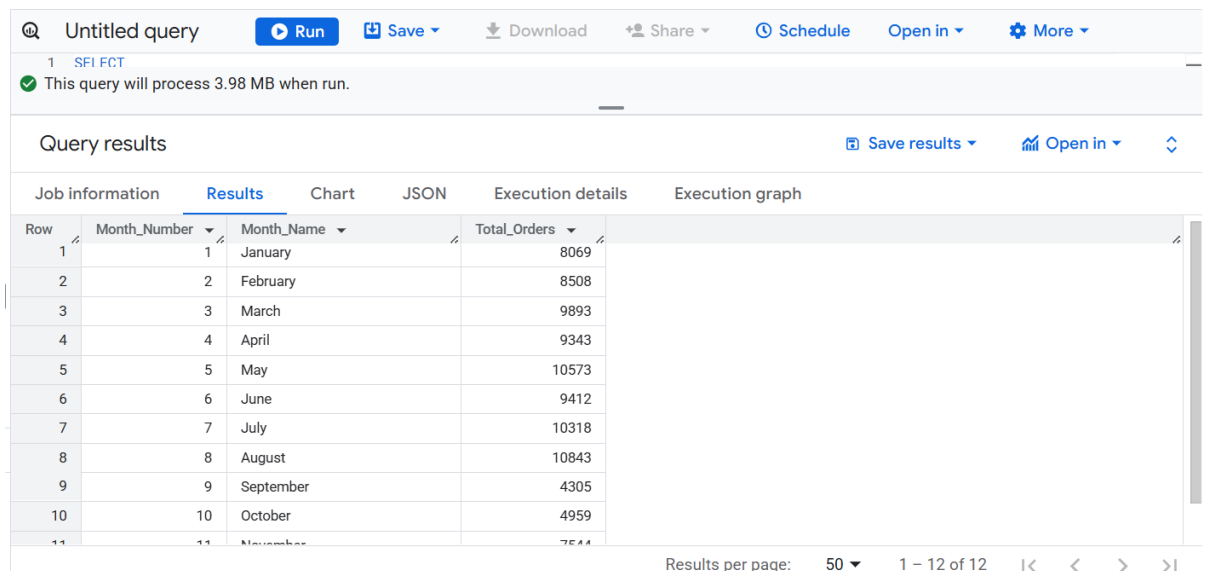
- **RECOMMENDATIONS –**

To maintain and further enhance sales trends, the company should focus on **regularly updating inventory**. Since many brands frequently launch new products, ensuring that these trending items are readily available can attract more customers. Staying ahead of the competition by offering the latest products will help sustain customer interest and drive sales growth.

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

Ans-

```
SELECT EXTRACT(MONTH FROM order_purchase_timestamp) AS Month_Number,
       FORMAT_DATE("%B", order_purchase_timestamp) AS Month_Name,
       COUNT(order_id) AS Total_Orders
FROM `Target.orders`
GROUP BY Month_Name, Month_Number
ORDER BY Month_Number;
```



Query results

Row	Month_Number	Month_Name	Total_Orders
1	1	January	8069
2	2	February	8508
3	3	March	9893
4	4	April	9343
5	5	May	10573
6	6	June	9412
7	7	July	10318
8	8	August	10843
9	9	September	4305
10	10	October	4959
11	11	November	7544

Results per page: 50 | 1 - 12 of 12

• INSIGHTS -

Analyzing monthly seasonality reveals that **May, July, and August** are **peak months** for customer orders, likely influenced by events such as **Black Friday** and other festivals when people stock up. In contrast, **September and October** record the **lowest sales**, possibly because customers still have ample stock from previous festive purchases.

• RECOMMENDATIONS -

➤ Peak Months:

- Ensure inventory is **well-stocked** to meet high demand.
- Consider **increasing staffing** to maintain a seamless shopping experience.

➤ Low-Sales Months:

- Introduce **special deals** like "**Buy 1 Get 1 Free**" or **free home delivery** to encourage purchases.

➤ Unique Strategy:

- During festivals, customers often buy items they don't end up using. To minimize waste and reduce inventory loss, introduce a **buy-back program** where customers can return unused items at **half price**. These items can then be **refurbished or resold**, allowing the company to recover value and reduce wastage.

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

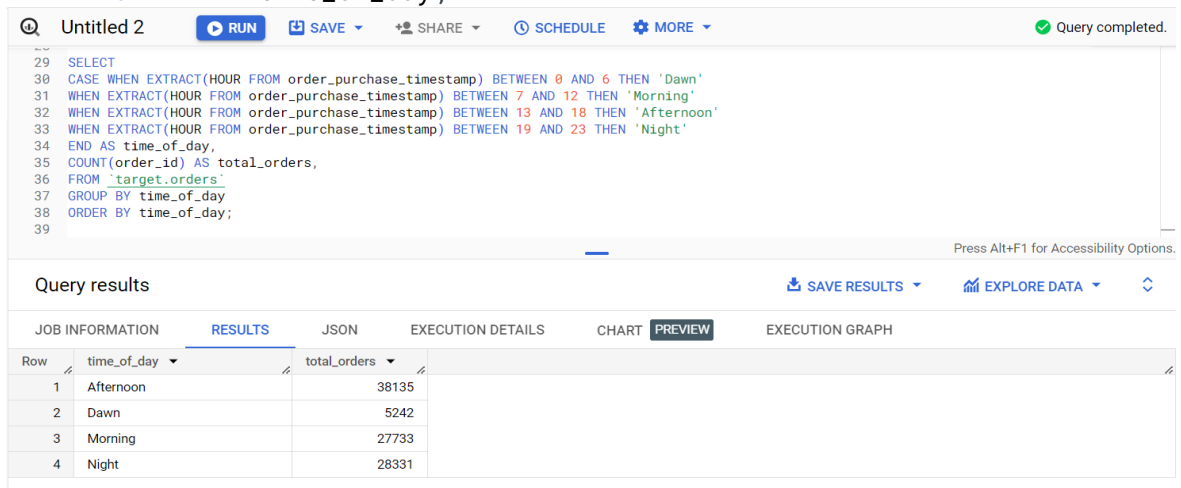
0-6 hrs : Dawn

7-12 hrs : Mornings

13-18 hrs : Afternoon

19-23 hrs : Night

Ans- `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_day, COUNT(order_id) AS total_orders,`
`FROM `target.orders``
`GROUP BY time_of_day`
`ORDER BY time_of_day;`



Query results

Row	time_of_day	total_orders
1	Afternoon	38135
2	Dawn	5242
3	Morning	27733
4	Night	28331

• INSIGHTS:

Analyzing order data by time reveals that customers place the **highest number of orders in the Afternoon**, followed by **Night, Morning, and Dawn**. This pattern highlights a preference for shopping during midday, with relatively lower activity during the early morning hours.

• RECOMMENDATIONS:

To make the most of customer behavior patterns, we can implement targeted strategies based on **peak** and **low-activity times**. During **peak hours**, especially in the **afternoon** when orders are highest, we can introduce **attractive offers** to further boost sales. For instance, during the **summer season**, offering a **free cold drink** with every order can draw more customers. On the other hand, during **low-activity times** like **dawn**, we can **reduce late-night charges** to encourage more purchases. Additionally, offering **quick delivery** of essential **medical products** during these early hours can cater to **emergency needs**, fostering **customer loyalty** and trust. By aligning our strategies with customer ordering patterns, we can maximize **engagement** and **revenue** throughout the day.

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

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

Ans-

```
SELECT CONCAT(EXTRACT(YEAR FROM order_purchase_timestamp), '-',
EXTRACT(MONTH FROM order_purchase_timestamp)) AS year_month,
c.customer_state, COUNT(*) AS order_numbers
FROM `target.orders` o JOIN `target.customers` c
ON o.customer_id = c.customer_id
GROUP BY year_month, c.customer_state
ORDER BY year_month, c.customer_state;
```

Query results

Row	year_month	customer_state	order_numbers
1	2016-10	AL	2
2	2016-10	BA	4
3	2016-10	CE	8
4	2016-10	DF	6
5	2016-10	ES	4
6	2016-10	GO	9
7	2016-10	MA	4
8	2016-10	MG	40
9	2016-10	MT	3
10	2016-10	PA	4

Results per page: 50 1 – 50 of 565

- **INSIGHTS -**

In October 2016, states like **MG (Minas Gerais)** and **MT (Mato Grosso)** had low order numbers, with **MG** at 40 orders and **MT** at 3. By August 2018, **SP (São Paulo)** saw a significant surge with 3,253 orders, followed by **RJ (Rio de Janeiro)** with 745 orders. States like **RS (Rio Grande do Sul)** and **SC (Santa Catarina)** showed consistent orders, while **PE (Pernambuco)** and **PB (Paraíba)** had fewer orders, especially in 2016.

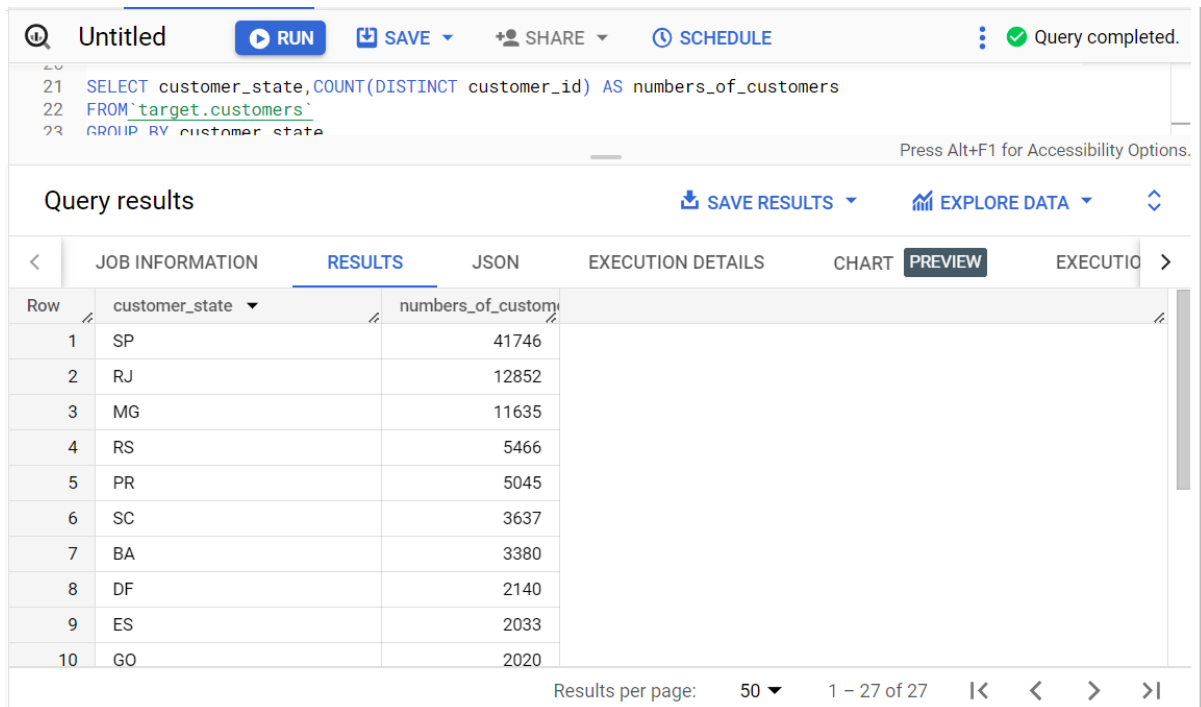
- **RECOMMENDATIONS –**

From the data, we observe that in **October 2016**, several states, such as **MG (Minas Gerais)** and **MT (Mato Grosso)**, contributed to the order numbers, with **MG** having 40 orders and **MT** having 3. As time progresses, the order volume increases significantly, particularly in **August 2018**, where **SP (São Paulo)** stands out with a total of **3,253 orders**, followed by **RJ (Rio de Janeiro)** with 745 orders. Other states such as **RS (Rio Grande do Sul)** and **SC (Santa Catarina)** show consistent order numbers in the later months. In contrast, some states, such as **PE (Pernambuco)** and **PB (Paraíba)**, show fewer orders, especially in earlier months like October 2016.

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

Ans-

```
SELECT customer_state,
COUNT(DISTINCT customer_id) AS numbers_of_customers
FROM `target.customers`
GROUP BY customer_state
ORDER BY numbers_of_customers DESC;
```



The screenshot shows a SQL query interface with a query editor at the top and a results table below. The query editor contains the following SQL code:

```
21 SELECT customer_state,COUNT(DISTINCT customer_id) AS numbers_of_customers
22 FROM `target.customers`
23 GROUP BY customer_state
```

The results table, titled "Query results", displays the following data:

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

The interface also includes a "Query completed." status message and a "Results per page: 50" indicator.

- **INSIGHTS:**

Our analysis of customer distribution across Brazilian states reveals that São Paulo (SP), Rio de Janeiro (RJ), and Minas Gerais (MG) are the top states with a significant customer base, while Acre (AC), Amapá (AP), and Roraima (RR) have fewer clients.

- **RECOMMENDATIONS:**

To increase the customer base in states with fewer clients like **Acre (AC)**, **Amapá (AP)**, and **Roraima (RR)**, we should conduct **targeted marketing campaigns** to boost brand awareness. Offering **special promotions** and **localized deals** can help attract new customers and build a stronger presence in these regions.

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).

Ans-

```
SELECT
ROUND((SUM(CASE WHEN EXTRACT(YEAR FROM
o.order_purchase_timestamp ) = 2018 AND EXTRACT(MONTH FROM
o.order_purchase_timestamp) <= 8 THEN p.payment_value ELSE 0
END) -
SUM(CASE WHEN EXTRACT(YEAR FROM o.order_purchase_timestamp) =
2017 AND EXTRACT(MONTH FROM o.order_purchase_timestamp) <= 8
THEN p.payment_value ELSE 0 END)) /
SUM(CASE WHEN EXTRACT(YEAR FROM o.order_purchase_timestamp) =
2017 AND EXTRACT(MONTH FROM o.order_purchase_timestamp) <= 8
THEN p.payment_value ELSE 1 END) * 100, 2) AS
percentage_increase
FROM `target.payments` AS p join `target.orders` as o
ON p.order_id = o.order_id
```

The screenshot shows a SQL query execution interface. At the top, there's a toolbar with buttons for 'RUN', 'SAVE', 'SHARE', and 'SCHEDULE'. A status bar indicates 'Query completed.' Below the toolbar, the SQL query is displayed in a text area. The query calculates the percentage increase in the cost of orders from 2017 to 2018, considering only orders from January to August. The query is as follows:

```
SELECT
ROUND((SUM(CASE WHEN EXTRACT(YEAR FROM o.order_purchase_timestamp ) = 2018 AND EXTRACT(MONTH FROM o.
order_purchase_timestamp) <= 8 THEN p.payment_value ELSE 0 END) -
SUM(CASE WHEN EXTRACT(YEAR FROM o.order_purchase_timestamp) = 2017 AND EXTRACT(MONTH FROM o.
order_purchase_timestamp) <= 8 THEN p.payment_value ELSE 0 END)) /
SUM(CASE WHEN EXTRACT(YEAR FROM o.order_purchase_timestamp) = 2017 AND EXTRACT(MONTH FROM o.
order_purchase_timestamp) <= 8 THEN p.payment_value ELSE 1 END) * 100, 2) AS percentage_increase
FROM `target.payments` AS p join `target.orders` as o
ON p.order_id = o.order_id
```

Below the query, the 'Query results' section is visible. It includes tabs for 'JOB INFORMATION', 'RESULTS', 'JSON', 'EXECUTION DETAILS', 'CHART', 'PREVIEW', and 'EXECUTION'. The 'RESULTS' tab is selected, showing a table with one row of results:

Row	percentage_increase
1	134.07

- INSIGHTS:**
 There was a **134.07% increase** in the cost of orders from **2017 to 2018**, indicating substantial growth in revenue.
- RECOMMENDATIONS:**
 Given this impressive increase, we should aim for a **160% growth** next year by focusing on **high-value product expansion**, **customer loyalty programs** and **targeted marketing strategies** to sustain the upward trend.

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

Ans-

```
SELECT customer_state,
ROUND(SUM(p.payment_value),2) AS total_order_price,
ROUND(AVG(p.payment_value),2) AS average_order_price
FROM `target.payments` AS p JOIN `target.orders` AS o
ON p.order_id = o.order_id
JOIN `target.customers` AS c
ON o.customer_id = c.customer_id
GROUP BY customer_state;
```

Query results

Row	customer_state	total_order_price	average_order_price
1	RJ	2144379.69	158.53
2	RS	890898.54	157.18
3	SP	5998226.96	137.5
4	DF	355141.08	161.13
5	PR	811156.38	154.15
6	MT	187029.29	195.23
7	MA	152523.02	198.86
8	AL	96962.06	227.08
9	MG	1872257.26	154.71
10	PE	324850.44	187.99

Results per page: 50 1 – 27 of 27

- INSIGHTS:**
 The **São Paulo (SP)** state shows **high sales** compared to other states. However, the **average price per order** is **significantly lower** than in other states.
- RECOMMENDATIONS:**
 To increase the **total order value**, we can experiment with offering **similar product pricing in other states** as in São Paulo (SP). This approach will help us determine whether **consistent pricing** across regions can boost overall sales.

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

Ans-

```
SELECT customer_state,
SUM(i.freight_value) AS total_freight_value,
AVG(i.freight_value) AS average_freight_value
FROM `target.order_items` AS i JOIN `target.orders` AS o
ON i.order_id = o.order_id
JOIN `target.customers` AS c
ON o.customer_id = c.customer_id
GROUP BY customer_state;
```

Query results

Row	customer_state	total_freight_value	average_freight_value
1	SP	718723.06999999...	15.14727539041...
2	RJ	305589.31000000...	20.96092393168...
3	PR	117851.68000000...	20.53165156794...
4	SC	89660.26000000...	21.47036877394...
5	DF	50625.49999999...	21.04135494596...
6	MG	270853.46000000...	20.63016680630...
7	PA	38699.30000000...	35.83268518518...
8	BA	100156.67999999...	26.36395893656...
9	GO	53114.97999999...	22.76681525932...
10	RS	135522.74000000...	21.73580433039...

Results per page: 50 1 - 27 of 27

- INSIGHTS:**

The **average freight value** is significantly higher in **RR (Roraima)** and **PB (Paraíba)**, possibly because these are **remote states**, leading to higher delivery costs. In contrast, **SP (São Paulo)** offers delivery at a **much lower freight value**.

- RECOMMENDATIONS:**

Since **SP** has **high overall sales**, we can consider **waiving the freight value** to encourage more purchases. For other states, **reducing freight costs** could help attract more customers and increase sales.

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

- 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.

Ans-

```
SELECT order_id,
TIMESTAMP_DIFF(order_delivered_customer_date,
order_purchase_timestamp, DAY) AS time_to_deliver,
TIMESTAMP_DIFF(order_estimated_delivery_date,
order_delivered_customer_date, DAY) AS diff_estimated_delivery
FROM `target.orders`;
```

Untitled	RUN	MORE	SAVE	SHARE	SCHEDULE	Query completed.
67	SELECT order_id,					
68	TIMESTAMP_DIFF(order_delivered_customer_date, order_purchase_timestamp, DAY) AS time_to_deliver,					
69	TIMESTAMP_DIFF(order_estimated_delivery_date, order_delivered_customer_date, DAY) AS diff_estimated_delivery					
						Press Alt+F1 for Accessibility Options.
Query results						
		SAVE RESULTS	EXPLORE DATA			
JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS	CHART	PREVIEW	EXECUTION GRAPH
Row	order_id	time_to_deliver	diff_estimated_delivery			
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			
Results per page: 50 1 - 50 of 99441						

- INSIGHTS -**

The **delivery time** for products ranges from **0 days to around 209 days**, indicating that some items are delivered **immediately**, while others take **over half a year** to reach customers. This inconsistency may impact **customer satisfaction**.

- RECOMMENDATIONS –**

To enhance the **delivery experience**, we should work on **reducing long delivery times** by possibly **hiring more delivery personnel** and optimizing **logistics processes**. Additionally, we must address discrepancies between **estimated and actual delivery times**. Identifying the **root causes of delays** will help us ensure that products arrive **on or before the promised date**, boosting **customer trust and satisfaction**.

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

Ans-

```
WITH StateFreight AS (
  SELECT customer_state, AVG(freight_value) AS avg_freight
  FROM `target.order_items` AS i JOIN `target.orders` AS o
  ON i.order_id = o.order_id
  JOIN `target.customers` AS c
  ON o.customer_id = c.customer_id
  GROUP BY customer_state
)
(SELECT customer_state, avg_freight
FROM StateFreight
ORDER BY avg_freight DESC
LIMIT 5)
UNION ALL
(SELECT customer_state, avg_freight
FROM StateFreight
ORDER BY avg_freight ASC
LIMIT 5);
```

Query results

Row	customer_state	avg_freight
1	RR	42.98442307692...
2	PB	42.72380398671...
3	RO	41.06971223021...
4	AC	40.07336956521...
5	PI	39.14797047970...
6	SP	15.14727539041...
7	PR	20.53165156794...
8	MG	20.63016680630...
9	RJ	20.96092393168...
10	DF	21.04135494596...

- INSIGHTS:**

States like **RR (Roraima)**, **PB (Paraíba)**, **RO (Rondônia)**, **AC (Acre)**, and **PI (Piauí)** have **high freight values**, while **SP (São Paulo)**, **PR (Paraná)**, **MG (Minas Gerais)**, **RJ (Rio de Janeiro)**, and **DF (Distrito Federal)** have **lower freight values**.

- RECOMMENDATIONS:**

To make products more **affordable** in states with **high freight costs** (RR, PB, RO, AC, PI), we should aim to **reduce freight charges** as much as possible. This strategy will make our products more **accessible** and help **attract more customers** from these regions. Focus efforts on bringing merchants of frequently purchased goods closer to the same zone or plan ahead to preserve storage space for these goods.

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

Ans-

```
WITH DeliveryTimeDays AS (
SELECT customer_state,
AVG(TIMESTAMP_DIFF(order_delivered_customer_date,
order_purchase_timestamp, DAY)) AS avg_delivery_time_days
FROM `target.orders` o JOIN `target.customers` AS c
ON o.customer_id = c.customer_id
GROUP BY customer_state )
(SELECT customer_state, avg_delivery_time_days
FROM DeliveryTimeDays
ORDER BY avg_delivery_time_days DESC
LIMIT 5)
UNION ALL
(SELECT customer_state, avg_delivery_time_days
FROM DeliveryTimeDays
ORDER BY avg_delivery_time_days ASC
LIMIT 5);
```

Untitled ▶ RUN ⚙️ MORE 💾 SAVE 👤 SHARE 🕒 SCHEDULE ✅ Query completed.

```
88 WITH DeliveryTimeDays AS (
89 SELECT customer_state,
90 AVG(TIMESTAMP_DIFF(order_delivered_customer_date, order_purchase_timestamp, DAY)) AS avg_delivery_time_days
91 FROM `target.orders` o JOIN `target.customers` AS c
```

Press Alt+F1 for Accessibility Options

Query results 📄 SAVE RESULTS 📊 EXPLORE DATA

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	CHART	PREVIEW	EXECUTION GRAPH
Row	customer_state	avg_delivery_time_days					
1	RR	28.97560975609...					
2	AP	26.73134328358...					
3	AM	25.98620689655...					
4	AL	24.04030226700...					
5	PA	23.31606765327...					
6	SP	8.298061489072...					
7	PR	11.52671135486...					
8	MG	11.54381329810...					
9	DF	12.50913461538...					
10	SC	14.47956019171...					

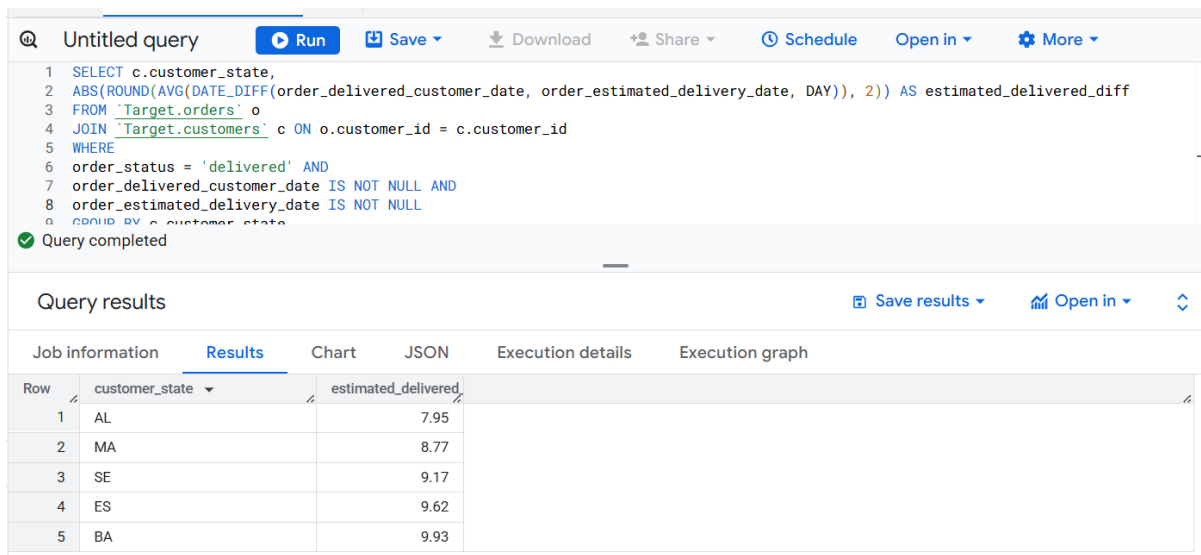
- INSIGHTS:**
 States like **SP (São Paulo)**, **PR (Paraná)**, **MG (Minas Gerais)**, **DF (Distrito Federal)**, and **SC (Santa Catarina)** have **shorter delivery times**, while **RR (Roraima)**, **AP (Amapá)**, **AM (Amazonas)**, **AL (Alagoas)**, and **PA (Pará)** experience **longer delivery times**.
- RECOMMENDATIONS:**
 To **reduce delivery times** in states with delays (RR, AP, AM, AL, PA), we can consider **hiring more delivery personnel** or **partnering with reliable delivery services**. This will help ensure **faster and more efficient deliveries**, leading to **better customer satisfaction**.

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

Ans-

```
SELECT c.customer_state,
ABS(ROUND(AVG(DATE_DIFF(order_delivered_customer_date,
order_estimated_delivery_date, DAY)), 2)) AS
estimated_delivered_diff
FROM `Target.orders` o
JOIN `Target.customers` c ON o.customer_id = c.customer_id
WHERE
order_status = 'delivered' AND
order_delivered_customer_date IS NOT NULL AND
order_estimated_delivery_date IS NOT NULL
GROUP BY c.customer_state
ORDER BY estimated_delivered_diff
LIMIT 5;
```



The screenshot shows a SQL query editor with a toolbar at the top containing icons for Run, Save, Download, Share, Schedule, Open in, and More. The query is pasted into the editor and has been executed successfully, as indicated by a green checkmark and the text 'Query completed'.

Below the query editor, the 'Query results' section is displayed. It includes tabs for Job information, Results (selected), Chart, JSON, Execution details, and Execution graph. The 'Results' tab shows a table with 5 rows of data.

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

- **INSIGHTS:**

The states with the **fastest order deliveries** compared to the **estimated delivery date** are **AL (Alagoas)**, **MA (Maranhão)**, **SE (Sergipe)**, **ES (Espírito Santo)**, and **BA (Bahia)**. These states have shown quicker delivery times, indicating efficient logistics and faster fulfillment.

- **RECOMMENDATIONS:**

To further enhance customer satisfaction, we should analyze the logistics strategies in these states and consider implementing similar approaches in regions with slower deliveries. This could help us **reduce delivery times** across the board and improve **overall service efficiency**.

6. Analysis based on the payments:

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

Ans-

```
SELECT
EXTRACT(YEAR FROM order_purchase_timestamp) AS year,
EXTRACT(MONTH FROM order_purchase_timestamp) AS month,
payment_type,
COUNT(DISTINCT o.order_id) AS num_orders
FROM `target.orders` o join `target.payments` p
ON o.order_id = p.order_id
GROUP BY year, month, payment_type
ORDER BY year, month, payment_type;
```

Untitled ▶ RUN ⚙️ MORE 💾 SAVE 👤 SHARE 🕒 SCHEDULE ✅ Query completed.

```
121 SELECT
122 EXTRACT(YEAR FROM order_purchase_timestamp) AS year,
123 EXTRACT(MONTH FROM order_purchase_timestamp) AS month,
```

Press Alt+F1 for Accessibility Options

Query results 📄 SAVE RESULTS 📊 EXPLORE DATA ⬆

JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS	CHART	PREVIEW	EXECUTION GRAPH
Row	year	month	payment_type	num_orders		
1	2016	9	credit_card	3		
2	2016	10	UPI	63		
3	2016	10	credit_card	253		
4	2016	10	debit_card	2		
5	2016	10	voucher	11		
6	2016	12	credit_card	1		
7	2017	1	UPI	197		
8	2017	1	credit_card	582		
9	2017	1	debit_card	9		
10	2017	1	voucher	33		

Results per page: 50 1 - 50 of 90 ⏪ ⏩

- **INSIGHTS:**

Credit card payments are the most dominant across months, especially in **July and August 2018**, with peaks of **4738** and **4963** orders. **UPI** saw significant use in early 2017 but tapered off, while **voucher** payments remained steady, peaking in **July and August 2018**. **Debit cards** had moderate use, and a small number of orders were recorded under "**not_defined**" payment type.

- **RECOMMENDATIONS:**

To capitalize on trends, focus on **credit card promotions** and consider offering **UPI incentives** to revive its use. Further, increase **voucher-based campaigns** during peak months and address the "**not_defined**" payment issue to improve data accuracy and customer experience.

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

Ans-

```
SELECT payment_installments,
COUNT(DISTINCT order_id) AS num_orders_placed
FROM `target.payments`
GROUP BY payment_installments
ORDER BY payment_installments;
```

The screenshot shows a SQL query execution interface. At the top, there's a toolbar with buttons for 'RUN', 'MORE', 'SAVE', 'SHARE', and 'SCHEDULE'. A status bar indicates 'Query completed.' Below the toolbar, the query text is displayed:

```
133 SELECT payment_installments,
134 COUNT(DISTINCT order_id) AS num_orders_placed
```

. The 'Query results' section is active, showing a table with two columns: 'payment_installment' and 'num_orders_placed'. The table has 10 rows of data. At the bottom, there's a pagination bar showing 'Results per page: 50' and '1 - 24 of 24'.

Row	payment_installment	num_orders_placed
1	0	2
2	1	49060
3	2	12389
4	3	10443
5	4	7088
6	5	5234
7	6	3916
8	7	1623
9	8	4253
10	9	644

- **INSIGHTS:**

The majority of orders are placed with **12 installments** (133 orders), followed by **15 installments** (74 orders). Higher installment plans like **23** and **24 installments** are rarely used, suggesting that customers prefer standard installment options.

- **RECOMMENDATIONS:**

To optimize sales, the company should focus on promoting the **12 installment** option, while also considering offering incentives to encourage customers to choose **higher installment plans**. Additionally, analyzing customer preferences for **lower installment options** can help create more flexible payment plans tailored to customer needs.