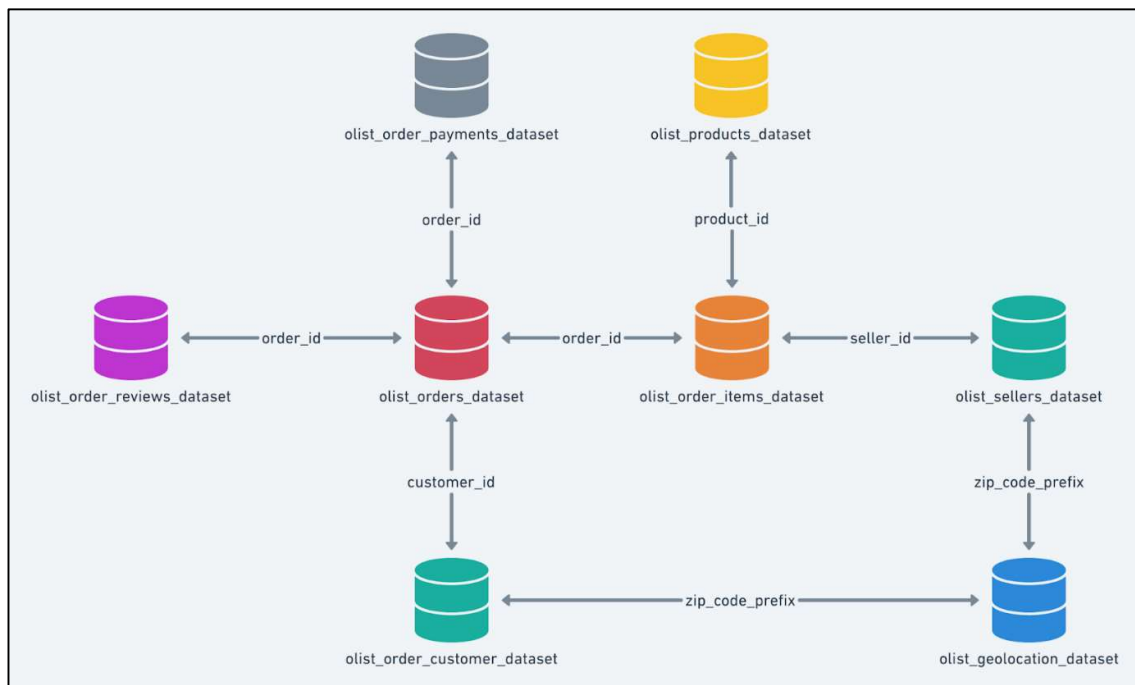


## Context of the Target Case Study:

Target is a globally renowned brand and a prominent retailer in the United States. Target makes itself a preferred shopping destination by offering outstanding value, inspiration, innovation and an exceptional guest experience that no other retailer can deliver.

This particular business case focuses on the operations of Target in Brazil and provides insightful information about 100,000 orders placed between 2016 and 2018. The dataset offers a comprehensive view of various dimensions including the order status, price, payment and freight performance, customer location, product attributes, and customer reviews.

By analyzing this extensive dataset, it becomes possible to gain valuable insights into Target's operations in Brazil. The information can shed light on various aspects of the business, such as order processing, pricing strategies, payment and shipping efficiency, customer demographics, product characteristics, and customer satisfaction levels.



**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.

Query:

```
SELECT column_name, data_type
FROM `plucky-plexus-396716.Target_bcs.INFORMATION_SCHEMA.COLUMNS`
WHERE table_name = 'customers';
```

Results:

customers

QUERY SHARE COPY SNAPSHOT DELETE EXPORT

SCHEMADETAILSPREVIEWLINEAGEDATA PROFILEDATA QUALITY

Filter

Enter property name or value

<input type="checkbox"/>	Field name	Type	Mode	Key	Collation	Default Value	Policy Tags	Description
<input type="checkbox"/>	customer_id	STRING	NULLABLE					
<input type="checkbox"/>	customer_unique_id	STRING	NULLABLE					
<input type="checkbox"/>	customer_zip_code_prefix	INTEGER	NULLABLE					
<input type="checkbox"/>	customer_city	STRING	NULLABLE					
<input type="checkbox"/>	customer_state	STRING	NULLABLE					

14  
15 select column\_name, data\_type  
16 from `plucky-plexus-396716.Target\_bcs.INFORMATION\_SCHEMA.COLUMNS`  
17 where table\_name = 'customers';

Query results

SAVE RESULTS

JOB INFORMATIONRESULTSJSONEXECUTION DETAILSCHARTPREVIEWEXECUTION GRAPH

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: To work on any data, we need to know the data before. Here, we have seen the data type of all the columns present in the customers table. We see that most of the customer data is stored as a string, meaning they are descriptive columns either storing name or address etc of a customer. The numeric fields such as zip codes, phone numbers will be saved as integers.

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

Query:

```
SELECT MIN(order_purchase_timestamp) DateRangea, MAX(order_purchase_timestamp)
DateRangeb
from `Target_bcs.orders`
```

Results:

1	select min(order_purchase_timestamp) DateRangea, max(order_purchase_timestamp) DateRangeb
2	from `Target_bcs.orders`

Query results	
JOB INFORMATION	RESULTS
Row	DateRangea
1	2016-09-04 21:15:19 UTC

JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS	CHART	PREVIEW	EXECUTION GRAPH
Row	DateRangea	DateRangeb				
1	2016-09-04 21:15:19 UTC	2018-10-17 17:30:18 UTC				

Given: We have taken order\_purchase\_timestamp in orders table as the time when orders were placed.

Insights: The date range queried above is the range between which the orders are placed at target by its customers across all given states. DateRangea is the date when first order was placed by any customer (2016 September 4), DateRangeb is the date when the most recent order was placed by any customer (2018 October 17).

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

Given: We have taken customer\_id from customers table as the customers who ordered during the said period.

Insights: We have calculated the total number of cities and states that customers belong to who have placed orders at Target. The highest number of customers belong to Sao Paulo (SP State) with 15540 customers. It has more customers than next 5 highest customer states combined. More than 1200 states have single customers who have placed orders. Either the population of such states are low or target has a limited marketing, sales and distribution in the area.

Query:

```
SELECT DISTINCT c.customer_city City, c.customer_state State, COUNT(o.customer_id)
Customer_Count
FROM `Target_bcs.customers` c JOIN `Target_bcs.orders` o ON c.customer_id =
o.customer_id
GROUP BY City, State
ORDER BY Customer_Count DESC;
```

## Results:

```
5 select distinct c.customer_city City, c.customer_state State, count(o.customer_id) Customer_Count
6 from `Target_bcs.customers` c join `Target_bcs.orders` o on c.customer_id = o.customer_id
7 group by City, State
8 order by Customer_Count DESC;
```

Query results

JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS	CHART	PREVIEW	EXECUTION GRAPH
Row	City	State	Customer_Count			
1	sao paulo	SP	15540			
2	rio de janeiro	RJ	6882			
3	belo horizonte	MG	2773			
4	brasilia	DF	2131			
5	curitiba	PR	1521			
6	campinas	SP	1444			
7	porto alegre	RS	1379			
8	salvador	BA	1245			
9	guarulhos	SP	1189			
10	sao bernardo do campo	SP	938			
11	niteroi	RJ	849			

## 2.In-depth Exploration:

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

YES.

Insights: From the given dataset, it seems that the orders each year are increasing. The second year of Target, the orders skyrocketed and reached more than 130 times the previous year. It shows that the Target was gaining fame and trust of customers over the years.

There is a growing trend in the number of orders by month as well over the past few years, with some fluctuations during few months in between. Overall, a growing trend of order count.

There might be an increase in population, increasing customer base, increasing fame of the company over the years through marketing and distribution.

## Query:

By Year:

```
WITH cte AS (SELECT
COUNT(order_id) AS order_count,
EXTRACT(year FROM order_purchase_timestamp) AS YearofPurchases
FROM `plucky-plexus-396716.Target_bcs.orders`
GROUP BY EXTRACT(year FROM order_purchase_timestamp))
```

```
SELECT YearofPurchases, order_count FROM cte
ORDER BY YearofPurchases;
```

Untitled

RUNSAVESHARESCHEDULEMORE

```
1 with cte as (select
2 count(order_id) as order_count,
3 extract(year from order_purchase_timestamp) as YearofPurchases
4 from `plucky-plexus-396716.Target_bcs.orders`
5 group by extract(year from order_purchase_timestamp))
6
7 select YearofPurchases, order_count from cte
8 order by YearofPurchases;
```

Processing location: asia-south2

Query results

SAVE

JOB INFORMATIONRESULTSJSONEXECUTION DETAILSCHARTPREVIEWEXECUTION GRAPH

Row	YearofPurchases	order_count
1	2016	329
2	2017	45101
3	2018	54011

By Month and year:

```
SELECT
  EXTRACT(YEAR FROM o.order_purchase_timestamp) AS year,
  EXTRACT(month FROM o.order_purchase_timestamp) AS month,
  FORMAT_TIMESTAMP('%B', TIMESTAMP(o.order_purchase_timestamp)) AS month_name,
  COUNT(DISTINCT o.order_id) AS order_count
FROM
  `plucky-plexus-396716.Target_bcs.orders` o
GROUP BY
  year, month, month_name
ORDER BY
  year, month;
```

Results:

Google Cloud | Scaler-DSML-SQL | Search (/) for resources, docs, products, and more | Search

Untitled | RUN | SAVE | SHARE | SCHEDULE | MORE | This script will

```

10 SELECT
11   EXTRACT(YEAR FROM o.order_purchase_timestamp) AS year,
12   EXTRACT(MONTH FROM o.order_purchase_timestamp) AS month,
13   format_timestamp('%B', timestamp(o.order_purchase_timestamp)) AS month_name,
14   COUNT(DISTINCT o.order_id) AS order_count
15 FROM
16   `plucky-plexus-396716.Target_bcs.orders` o
17 GROUP BY
18   year, month, month_name
19 ORDER BY
20   year, month:

```

Query results | SAVE RESULTS | E

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	CHART	PREVIEW	EXECUTION GRAPH
Row	year	month	month_name	order_count			
1	2016	9	September	4			
2	2016	10	October	324			
3	2016	12	December	1			
4	2017	1	January	800			
5	2017	2	February	1780			
6	2017	3	March	2682			
7	2017	4	April	2404			
8	2017	5	May	3700			
9	2017	6	June	3245			
10	2017	7	July	4026			
11	2017	8	August	4331			

Results per page: 50 | 1 - 25 of 25

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

Insights: We can see that the highest number of orders were placed during May, July and August. Usually, higher number of orders are placed in first half of the year, compared to the second half of the year (specially, last 4 months, the holiday season). Other than the above analysis, no other visible seasonality is seen per the given dataset.

Query:

```

SELECT
  EXTRACT(MONTH FROM order_purchase_timestamp) AS month,
  COUNT(DISTINCT order_id) AS order_count
FROM
  `plucky-plexus-396716.Target_bcs.orders`
GROUP BY month
ORDER BY month;

```

Results:

Google Cloud Scaler-DSML-SQL Search (/) for resources, docs, products, and more

customers geolocation order\_items order\_reviews orders

Untitled RUN SAVE SHARE SCHEDULE MORE

```

22 SELECT
23   EXTRACT(MONTH FROM order_purchase_timestamp) AS month,
24   COUNT(distinct order_id) AS order_count
25 FROM
26   `plucky-plexus-396716.Target_bcs.orders`
27 GROUP BY month
28 ORDER BY month;
29

```

Query results

JOB INFORMATION RESULTS JSON EXECUTION DETAILS CHART PREVIEW EXECUTION GRAPH

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

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

```

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 'Mornings'
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 hour,
COUNT(DISTINCT order_id) AS order_count
FROM `plucky-plexus-396716.Target_bcs.orders`
GROUP BY hour
ORDER BY order_count DESC;

```

## Results:

Untitled					
<pre>29 SELECT 30 case 31 when EXTRACT(hour FROM order_purchase_timestamp) between 0 and 6 then 'Dawn' 32 when EXTRACT(hour FROM order_purchase_timestamp) between 7 and 12 then 'Mornings' 33 when EXTRACT(hour FROM order_purchase_timestamp) between 13 and 18 then 'Afternoon' 34 when EXTRACT(hour FROM order_purchase_timestamp) between 19 and 23 then 'Night' 35 end as hour, 36 COUNT(DISTINCT order_id) AS order_count 37 FROM `plucky-plexus-396716.Target_bcs.orders` 38 GROUP BY hour 39 ORDER BY order_count desc;</pre>					
Query results					
JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	CHART PREVIEW EXECUTION GRAPH
Row	hour	order_count			
1	Afternoon	38135			
2	Night	28331			
3	Mornings	27733			
4	Dawn	5242			

Insights: Looking at the query results, we can say that the customers order at target the highest during afternoon, and night. Customer might get some time after their important work during afternoon; they even order online when they have some time at night as well. We assume that the timestamp recorded in the database is accurate.

Recommendation: This helps target to improve their marketing and advertising strategy to gain better number of consumers or get more and more orders from existing customers by providing them timely offers or discounts.

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

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

#### Query:

```
SELECT c.customer_state,
EXTRACT(month FROM order_purchase_timestamp) AS month,
FORMAT_DATE('%B',TIMESTAMP(order_purchase_timestamp)) AS month_name,
COUNT(DISTINCT order_id) AS order_count
FROM `plucky-plexus-396716.Target_bcs.orders` o
JOIN `plucky-plexus-396716.Target_bcs.customers` c
ON o.customer_id = c.customer_id
```



```
GROUP BY c.customer_state, month, month_name
ORDER BY c.customer_state, month;
```

Results:

The screenshot shows the Google Cloud BigQuery console. At the top, there's a search bar and a list of open queries. The main area displays a SQL query that extracts the month from the order purchase timestamp and counts the distinct order IDs for each month, grouped by customer state. Below the query editor, the 'Query results' section is active, showing a table with 12 rows and 5 columns: Row, customer\_state, month, month\_name, and order\_count. The results show that for each state (AC, SC, NC, GA, FL, AL, MS, TN, KY, WV, OH, PA), the number of orders varies by month, with January consistently having the highest count (8 orders) and December having the lowest (5 orders).

Row	customer_state	month	month_name	order_count
1	AC	1	January	8
2	AC	2	February	6
3	AC	3	March	4
4	AC	4	April	9
5	AC	5	May	10
6	AC	6	June	7
7	AC	7	July	9
8	AC	8	August	7
9	AC	9	September	5
10	AC	10	October	6
11	AC	11	November	5
12	AC	12	December	5

Insights: The query gives the order split between different months for all the states. We see that the count of orders for each month is greater for SP state, than any other state. There is no specific trend seen comparing the months as such.

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

Assumptions: We have taken the customer\_id column data from customers table as customers who are associated with Target.

Query:

```
SELECT c.customer_state,
COUNT(DISTINCT c.customer_id) AS No_of_customers
FROM `plucky-plexus-396716.Target_bcs.orders` o
JOIN `plucky-plexus-396716.Target_bcs.customers` c
ON o.customer_id = c.customer_id
GROUP BY c.customer_state
ORDER BY No_of_customers DESC, c.customer_state;
```

Results:

Customers

Geolocation

Order Items

Order Reviews

Orders

Payments

Products

Sellers

Untitled

Run

Save

Share

Schedule

More

```
52 SELECT c.customer_state,
53 COUNT(DISTINCT c.customer_id) AS No_of_customers
54 FROM `plucky-plexus-396716.Target_bcs.orders` o
55 JOIN `plucky-plexus-396716.Target_bcs.customers` c
56 ON o.customer_id = c.customer_id
57 GROUP BY c.customer_state
58 ORDER BY No_of_customers desc, c.customer_state;
59
```

Query results

Save Results

Export

Job Information

Results

JSON

Execution Details

Chart

Preview

Execution Graph

Row	customer_state	No_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
11	PE	1652
12	CE	1336
13	PA	975

Results per page: 50

1 - 27 of 27

Insights: Given query shows Distribution of customers across all the states. The highest number of customers (~41000) are located in SP State, which is why the orders are also the maximum in the SP State as seen in one of the above analyses. The customer count of RJ and MG are the next highest in order, matching with the ratio of orders as well. The lowest number of customers are from RR.

Recommendation: Target should increase their reach through marketing and advertising to the states where there are no or very low number of customers currently to gain wider customer base than before.

#### 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). You can use the "payment\_value" column in the payments table to get the cost of orders.

Query:

```

SELECT
  EXTRACT(MONTH FROM o.order_purchase_timestamp) AS month,
  FORMAT_DATE('%B',TIMESTAMP(o.order_purchase_timestamp)) AS month_name,
  ROUND(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),2) AS Cost2017,
  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),2) AS Cost2018,
  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

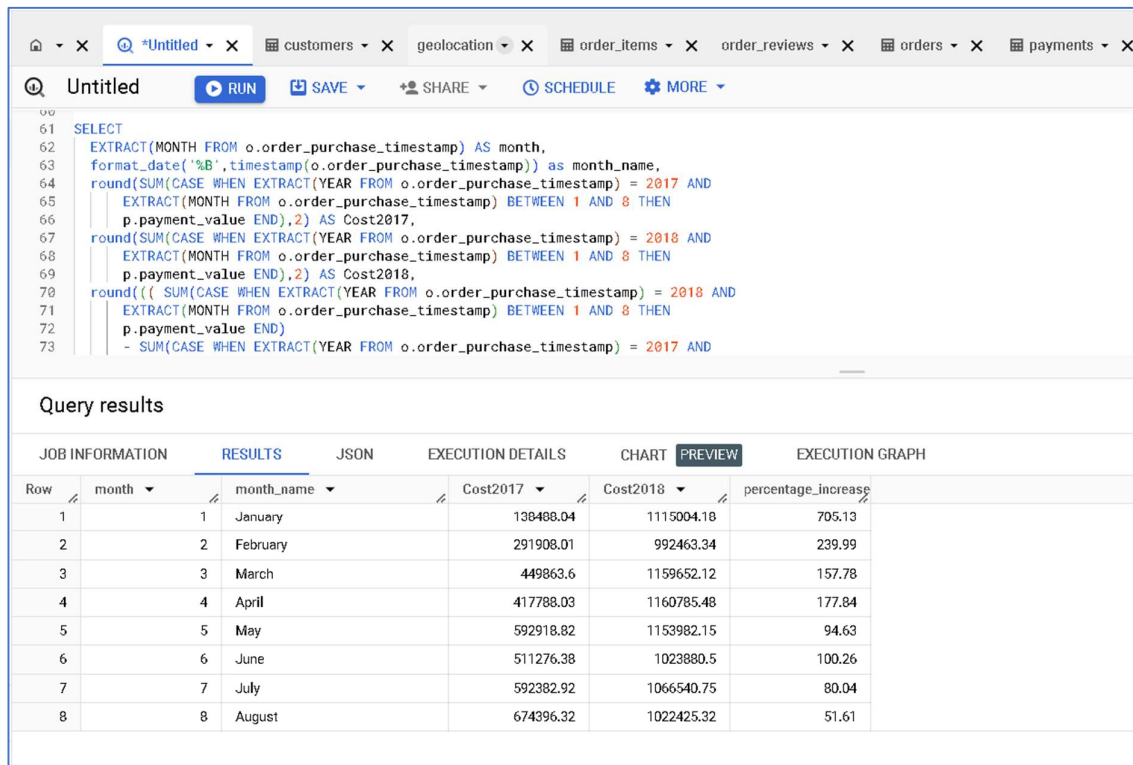
```

```

        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 percentage_increase
FROM `plucky-plexus-396716.Target_bcs.orders` o
JOIN `plucky-plexus-396716.Target_bcs.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 1,2 ORDER BY 1;

```

### Results:



The screenshot shows a SQL query editor with a query that calculates the percentage increase in order costs from 2017 to 2018, grouped by month. The results are displayed in a table with columns for Row, month, month\_name, Cost2017, Cost2018, and percentage\_increase.

Row	month	month_name	Cost2017	Cost2018	percentage_increase
1	1	January	130488.04	1115004.18	705.13
2	2	February	291908.01	992463.34	239.99
3	3	March	449863.6	1159652.12	157.78
4	4	April	417788.03	1160785.48	177.84
5	5	May	592918.82	1153982.15	94.63
6	6	June	511276.38	1023880.5	100.26
7	7	July	592382.92	1066540.75	80.04
8	8	August	674396.32	1022425.32	51.61

Insights: From the query results, we can conclude that the % increase in cost of orders from 2017 to 2018 is huge for each month. The greatest % increase was from Jan 2017 to Jan 2018, a massive 705% increase in cost of orders. Then the highest is in Feb and April 2017 to 2018. The increase in cost of orders from previous year to the next is a good sign for target and its business growth.

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

### Query:

```

SELECT c.customer_state, ROUND(SUM(oi.price),2) AS Order_Total,
ROUND(AVG(oi.price),2) AS Order_average

```

```

FROM `plucky-plexus-396716.Target_bcs.orders` o
JOIN `plucky-plexus-396716.Target_bcs.order_items` oi ON o.order_id = oi.order_id
JOIN `plucky-plexus-396716.Target_bcs.customers` c ON c.customer_id =
o.customer_id
GROUP BY c.customer_state
ORDER BY Order_Total desc, Order_average desc;

```

## Results:

Query results

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

Results per page: 50 1 - 27 of 27

Insights: The query results shows the total order and average order amount for each state till now at Target. The highest order total amount is from the most populated city, Sao Paulo, SP State. But the average amount of order is the lowest there. It means that more number of consumers are ordering , but each order is of a small order amount considering all the orders at Target. The next highest order amount is from RJ and MG, these 3 states have greatest order total amount, comparing the following states.

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

## Query:

```

SELECT c.customer_state, ROUND(SUM(oi.freight_value),2) AS frieght_Total,
ROUND(AVG(oi.freight_value),2) AS frieght_average
FROM `plucky-plexus-396716.Target_bcs.orders` o
JOIN `plucky-plexus-396716.Target_bcs.order_items` oi ON o.order_id = oi.order_id
JOIN `plucky-plexus-396716.Target_bcs.customers` c ON c.customer_id =
o.customer_id
GROUP BY c.customer_state
ORDER BY 2 DESC, 3 DESC;

```

## Results:

Untitled			
<pre> 92 select c.customer_state, round(sum(oi.freight_value),2) as freight_Total, round(avg(oi.freight_value),2) as freight_average 93 FROM 'plucky-plexus-396716.Target_bcs.orders' o 94 JOIN 'plucky-plexus-396716.Target_bcs.order_items' oi ON o.order_id = oi.order_id 95 JOIN 'plucky-plexus-396716.Target_bcs.customers' c ON c.customer_id = o.customer_id 96 group by c.customer_state 97 order by 2 desc, 3 desc; 98 </pre>			
Query results			
<div> JOB INFORMATION RESULTS JSON EXECUTION DETAILS CHART PREVIEW EXECUTION GRAPH </div>			
Row	customer_state	freight_Total	freight_average
1	SP	718723.07	15.15
2	RJ	305589.31	20.96
3	MG	270853.46	20.63
4	RS	135522.74	21.74
5	PR	117851.68	20.53
6	BA	100156.68	26.36
7	SC	89660.26	21.47
8	PE	59449.66	32.92
9	GO	53114.98	22.77
10	DF	50625.5	21.04
11	ES	49764.6	22.06
Results per page: 50 1 - 27 of 27			

Insights and Recommendations: The highest transport or shipping charges, called as the freight charges are for SP state, but the freight average is the lowest among all other states. Understanding the freight charges across all locations and coming up with a better strategy to build warehouses per locations having more orders or more freight charges can be good both to the consumers and company in the long run.

## 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.

You can calculate the delivery time and the difference between the estimated & actual delivery date using the given formula:

- **time\_to\_deliver** = order\_delivered\_customer\_date - order\_purchase\_timestamp
- **diff\_estimated\_delivery** = order\_estimated\_delivery\_date - order\_delivered\_customer\_date

Query:

```

SELECT order_id,
DATE_DIFF(order_delivered_customer_date, order_purchase_timestamp, day) AS
time_to_deliver,
DATE_DIFF(order_estimated_delivery_date, order_purchase_timestamp, DAY) AS
estimated_delivery_in_days,
DATE_DIFF(order_estimated_delivery_date, order_delivered_customer_date, day) AS
diff_estimated_delivery,

```

```

CASE WHEN DATE_DIFF(order_estimated_delivery_date,
order_delivered_customer_date,day) < 0 THEN 'Late Delivery'
      WHEN DATE_DIFF(order_estimated_delivery_date,
order_delivered_customer_date,day) =0 THEN 'On time Delivery'
      ELSE 'Early Delivery' END AS Delivered_status
FROM `plucky-plexus-396716.Target_bcs.orders`
WHERE DATE_DIFF(order_delivered_customer_date,order_purchase_timestamp,day) IS NOT
NULL AND DATE_DIFF(order_estimated_delivery_date,
order_delivered_customer_date,day) IS NOT NULL
ORDER BY time_to_deliver;

```

## Results:

Query results

Row	order_id	time_to_deliver	estimated_delivery_in_days	diff_estimated_delivery	Delivered_status
1	e65f1eeee1f52024ad1dc0d344...	0	10	9	Early Delivery
2	bb5a519e352b49b714192a02ff...	0	26	25	Early Delivery
3	434ceceee7d1a65fc65358a632b...	0	20	19	Early Delivery
4	d3ca7b82c922817b06e5ca2116...	0	12	11	Early Delivery
5	1d893dd7ca9f77ebf5f5f9f0201...	0	10	10	Early Delivery
6	d5f8eedc85190ba8858046f82d...	0	8	7	Early Delivery
7	79e324907160c8ea526fd8b943...	0	9	8	Early Delivery
8	38c1e3d4edea13cd0cf612d4cd...	0	17	16	Early Delivery
9	8339a608be0d84fca9d8a68b5...	0	28	27	Early Delivery
10	f349c0b62f69c3fae5c4d7d3f3a...	0	13	12	Early Delivery
11	f3c6775ba3d2d9fe2826f93b71f...	0	12	11	Early Delivery

Insights: Time taken to deliver the orders is a crucial part for every business. The earlier the business delivers, the better for it to gain customer trust and survive. The values of time\_to\_deliver, estimated\_delivery\_in\_days, diff\_esitimated\_delivery are in days. When the orders are delivered before estimated delivery date, it is considered to be an early delivery. Target has delivered many orders as early as within a day or as late as 209 days.

Recommendations: target needs to work on shipping and delivering the orders with intent to deliver it earlier or on time so that customer retention and accrual is easier.

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

## Query:

```

WITH cte AS (SELECT customer_state,avg_freight_value,
DENSE_RANK() OVER(ORDER BY avg_freight_value DESC) AS top_5,
DENSE_RANK() OVER(ORDER BY avg_freight_value) AS bottom_5 FROM
(SELECT c.customer_state AS customer_state,
ROUND(AVG(oi.freight_value), 2) AS avg_freight_value
FROM `plucky-plexus-396716.Target_bcs.order_items` oi
JOIN `plucky-plexus-396716.Target_bcs.orders` o ON oi.order_id = o.order_id

```

```

JOIN `plucky-plexus-396716.Target_bcs.customers` c ON o.customer_id = c.customer_id
GROUP BY c.customer_state) AS tbl1)
SELECT customer_state,avg_freight_value,'top 5 states' AS ranking FROM cte WHERE
top_5 <=5
UNION ALL
SELECT customer_state,avg_freight_value, 'bottom_5_states' AS ranking FROM cte
WHERE bottom_5 <=5
ORDER BY avg_freight_value DESC;

```

Results:

<div> <div> <div>Untitled</div> <div> <div>RUN</div> <div>SAVE</div> <div>SHARE</div> <div>SCHEDULE</div> <div>MORE</div> </div> </div> <div> <div>111</div> <div>with cte as (select customer_state,avg_freight_value,</div> <div>112</div> <div>dense_rank() over(order by avg_freight_value desc) as top_5,</div> <div>113</div> <div>dense_rank() over(order by avg_freight_value) as bottom_5 from</div> <div>114</div> <div>(select c.customer_state as customer_state,</div> <div>115</div> <div>ROUND(AVG(oi.freight_value), 2) AS avg_freight_value</div> <div>116</div> <div>from `plucky-plexus-396716.Target_bcs.order_items` oi</div> <div>117</div> <div>join `plucky-plexus-396716.Target_bcs.orders` o on oi.order_id = o.order_id</div> <div>118</div> <div>join `plucky-plexus-396716.Target_bcs.customers` c on o.customer_id = c.customer_id</div> <div>119</div> <div>group by c.customer_state) as tbl1)</div> <div>120</div> <div>select customer_state,avg_freight_value,'top 5 states' as ranking from cte where top_5 &lt;=5</div> <div>121</div> <div>union all</div> <div>122</div> <div>select customer_state,avg_freight_value,'bottom_5_states' as ranking from cte where bottom_5 &lt;=5</div> <div>123</div> <div>order by avg_freight_value desc;</div> </div> </div>					
Query results					
JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	CHART PREVIEW EXECUTION GRAPH
Row	customer_state	avg_freight_value	ranking		
1	RR	42.98	top 5 states		
2	PB	42.72	top 5 states		
3	RO	41.07	top 5 states		
4	AC	40.07	top 5 states		
5	PI	39.15	top 5 states		
6	DF	21.04	bottom_5_states		
7	RJ	20.96	bottom_5_states		
8	MG	20.63	bottom_5_states		
9	PR	20.53	bottom_5_states		
10	SP	15.15	bottom_5_states		

Insights: SP has the lowest average freight value, and RR has the highest. The query shows the top 5 and bottom 5 states in terms of the average freight value.

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

Query:

```

WITH cte AS (SELECT customer_state,average_time_to_deliver,
DENSE_RANK() OVER(ORDER BY average_time_to_deliver) AS top_5,
DENSE_RANK() OVER(ORDER BY average_time_to_deliver DESC) AS bottom_5 FROM
(SELECT c.customer_state AS customer_state,
ROUND(AVG(DATE_DIFF(o.order_delivered_customer_date,o.order_purchase_timestamp,day)
),2) AS average_time_to_deliver
FROM `plucky-plexus-396716.Target_bcs.orders` o
JOIN `plucky-plexus-396716.Target_bcs.customers` c ON o.customer_id = c.customer_id
GROUP BY c.customer_state) AS tbl1)

```





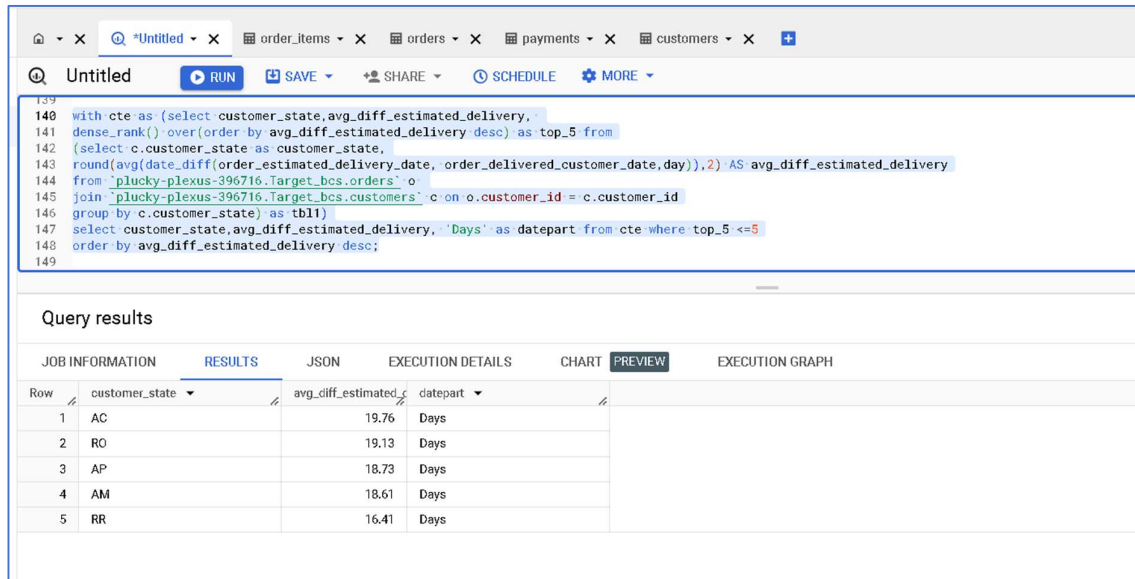


```

ROUND(AVG(DATE_DIFF(order_estimated_delivery_date,
order_delivered_customer_date,day)),2) AS avg_diff_estimated_delivery
FROM `plucky-plexus-396716.Target_bcs.orders` o
JOIN `plucky-plexus-396716.Target_bcs.customers` c ON o.customer_id = c.customer_id
GROUP BY c.customer_state) AS tbl1)
SELECT customer_state,avg_diff_estimated_delivery, 'Days' AS datepart FROM cte
WHERE top_5 <=5
ORDER BY avg_diff_estimated_delivery DESC;

```

### Results:



Query results

Row	customer_state	avg_diff_estimated_delivery	datepart
1	AC	19.76	Days
2	RO	19.13	Days
3	AP	18.73	Days
4	AM	18.61	Days
5	RR	16.41	Days

**Insights:** The query results show that for these states mentioned, the average time to deliver is way faster than the expected/ previously estimated delivery date. The shipping process or the delivery process was better than the customer expected in the top 5 fast delivery states. This increases customer satisfaction.

**Recommendations:** Target should check on the states where difference between actual delivery and expected delivery date is very less or the actual delivery date exceeds the estimated delivery, it can affect the business in negative way if they increase.

## 6. Analysis based on the payments:

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

### Query:

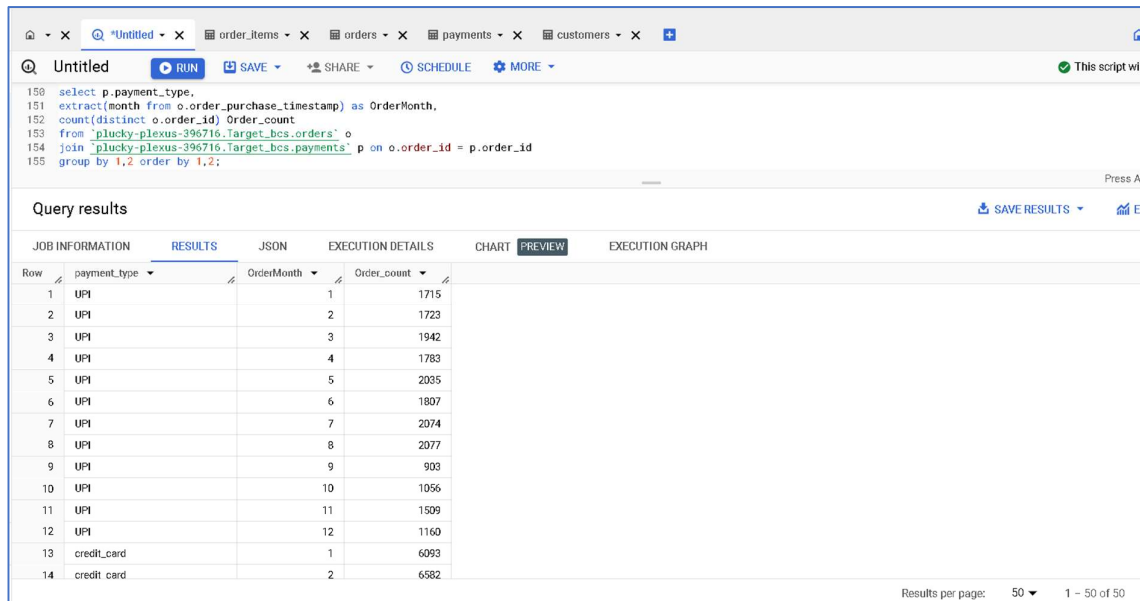
```

SELECT
p.payment_type,
EXTRACT(month FROM o.order_purchase_timestamp) AS OrderMonth,
COUNT(DISTINCT o.order_id) Order_count
FROM `plucky-plexus-396716.Target_bcs.orders` o
JOIN `plucky-plexus-396716.Target_bcs.payments` p ON o.order_id = p.order_id
GROUP BY 1,2

```

ORDER BY 1,2;

Results:



The screenshot shows a SQL query editor with a query and its results. The query is as follows:

```
150 select p.payment_type,
151       extract(month from o.order_purchase_timestamp) as OrderMonth,
152       count(distinct o.order_id) Order_count
153 from `plucky-plexus-396716.Target_bcs.orders` o
154 join `plucky-plexus-396716.Target_bcs.payments` p on o.order_id = p.order_id
155 group by 1,2 order by 1,2;
```

The results are displayed in a table with the following columns: Row, payment\_type, OrderMonth, and Order\_count. The results are as follows:

Row	payment_type	OrderMonth	Order_count
1	UPI	1	1715
2	UPI	2	1723
3	UPI	3	1942
4	UPI	4	1783
5	UPI	5	2035
6	UPI	6	1807
7	UPI	7	2074
8	UPI	8	2077
9	UPI	9	903
10	UPI	10	1056
11	UPI	11	1509
12	UPI	12	1160
13	credit_card	1	6093
14	credit card	2	6582

Insights: The results show that users pay with credit card the most while placing orders, debit card the least. UPI mode is the second highest, meaning the consumer base is technologically strong, and prefer better offers and discounts. Vouchers are used by some users, not really frequent.

Recommendations: Target can tie up with multiple banks and credit card companies to make the most out of the situation, to provide hefty offers and discounts on certain cards and payment modes, so that customer easily turns into heavy purchaser.

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

Query:

```
SELECT p.payment_installments,
COUNT(DISTINCT o.order_id) Order_count
FROM `plucky-plexus-396716.Target_bcs.orders` o
JOIN `plucky-plexus-396716.Target_bcs.payments` p ON o.order_id = p.order_id
WHERE o.order_status != 'canceled'
GROUP BY p.payment_installments ORDER BY p.payment_installments;
```

Results:

order\_itemsorderspaymentscustomers

Untitled

RUNSAVESHARESCHEDULEMORE

```
158 select p.payment_installments,
159 count(distinct o.order_id) Order_count
160 from 'plucky-plexus-396716.target_bcs.orders' o
161 join 'plucky-plexus-396716.target_bcs.payments' p on o.order_id = p.order_id
162 where o.order_status != 'canceled'
163 group by p.payment_installments order by p.payment_installments;
```

Query results

SAVE RESULTS

EXECUTION DETAILSCHARTPREVIEWEXECUTION GRAPH

JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS	CHART	PREVIEW	EXECUTION GRAPH
Row	payment_installment	Order_count				
1	0	2				
2	1	48732				
3	2	12329				
4	3	10374				
5	4	7046				
6	5	5204				
7	6	3894				
8	7	1617				
9	8	4224				
10	9	638				
11	10	5279				
12	11	22				
13	12	133				
14	13	133				
15	14	133				
16	15	133				
17	16	133				
18	17	133				
19	18	133				
20	19	133				
21	20	133				
22	21	133				
23	22	133				
24	23	133				
25	24	133				
26	25	133				
27	26	133				
28	27	133				
29	28	133				
30	29	133				
31	30	133				
32	31	133				
33	32	133				
34	33	133				
35	34	133				
36	35	133				
37	36	133				
38	37	133				
39	38	133				
40	39	133				
41	40	133				
42	41	133				
43	42	133				
44	43	133				
45	44	133				
46	45	133				
47	46	133				
48	47	133				
49	48	133				
50	49	133				
51	50	133				
52	51	133				
53	52	133				
54	53	133				
55	54	133				
56	55	133				
57	56	133				
58	57	133				
59	58	133				
60	59	133				
61	60	133				
62	61	133				
63	62	133				
64	63	133				
65	64	133				
66	65	133				
67	66	133				
68	67	133				
69	68	133				
70	69	133				
71	70	133				
72	71	133				
73	72	133				
74	73	133				
75	74	133				
76	75	133				
77	76	133				
78	77	133				
79	78	133				
80	79	133				
81	80	133				
82	81	133				
83	82	133				
84	83	133				
85	84	133				
86	85	133				
87	86	133				
88	87	133				
89	88	133				
90	89	133				
91	90	133				
92	91	133				
93	92	133				
94	93	133				
95	94	133				
96	95	133				
97	96	133				
98	97	133				
99	98	133				
100	99	133				
101	100	133				
102	101	133				
103	102	133				
104	103	133				
105	104	133				
106	105	133				
107	106	133				
108	107	133				
109	108	133				
110	109	133				
111	110	133				
112	111	133				
113	112	133				
114	113	133				
115	114	133				
116	115	133				
117	116	133				
118	117	133				
119	118	133				
120	119	133				
121	120	133				
122	121	133				
123	122	133				
124	123	133				
125	124	133				
126	125	133				
127	126	133				
128	127	133				
129	128	133				
130	129	133				
131	130	133				
132	131	133				
133	132	133				
134	133	133				
135	134	133				
136	135	133				
137	136	133				
138	137	133				
139	138	133				
140	139	133				
141	140	133				
142	141	133				
143	142	133				
144	143	133				
145	144	133				
146	145	133				
147	146	133				
148	147	133				
149	148	133				
150	149	133				
151	150	133				
152	151	133				
153	152	133				
154	153	133				
155	154	133				
156	155	133				
157	156	133				
158	157	133				
159	158	133				
160	159	133				
161	160	133				
162	161	133				
163	162	133				
164	163	133				
165	164	133				
166	165	133				
167	166	133				
168	167	133				
169	168	133				
170	169	133				
171	170	133				
172	171	133				
173	172	133				
174	173	133				
175	174	133				
176	175	133				
177	176	133				
178	177	133				
179	178	133				
180	179	133				
181	180	133				
182	181	133				
183	182	133				
184	183	133				
185	184	133				
186	185	133				
187	186	133				
188	187	133				
189	188	133				
190	189	133				
191	190	133				
192	191	133				
193	192	133				
194	193	133				
195	194	133				
196	195	133				
197	196	133				
198	197	133				
199	198	133				
200	199	133				

Results per page: 501 - 24 of 24