

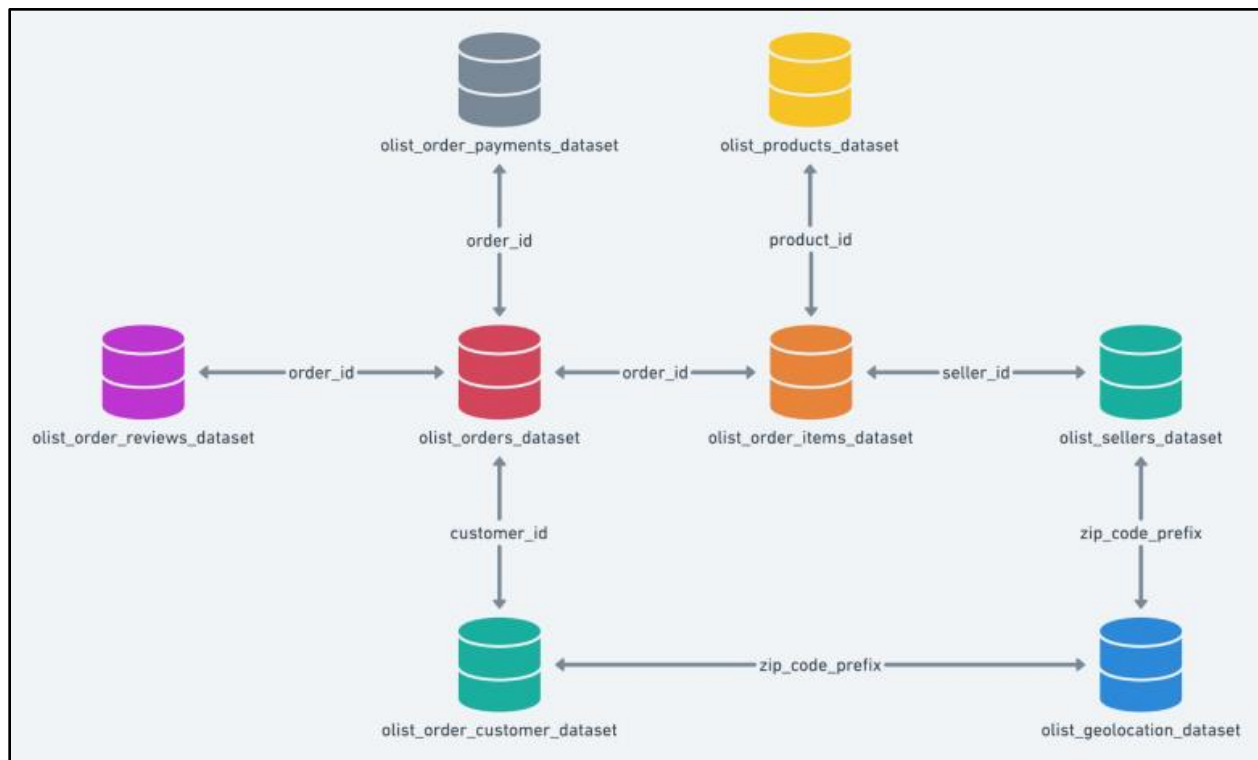


# TARGET SQL – BUSINESS CASE

BY - AMOL SUSHIR

**From company's perspective:**

- 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 business case focuses on the operations of Target in Brazil and provides insightful information about 100,000 orders placed between 2016 and 2018. The dataset offers a comprehensive view of various dimensions including the order status, price, payment and freight performance, customer location, product attributes, and customer reviews.
- By analyzing this extensive dataset, it becomes possible to gain valuable insights into Target's operations in Brazil. The information can shed light on various aspects of the business, such as order processing, pricing strategies, payment and shipping efficiency, customer demographics, product characteristics, and customer satisfaction levels.

**Dataset schema:**

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

A. Data type of all columns in the “customers” table.

Ans:

```
SELECT
  column_name,
  data_type
FROM target.INFORMATION_SCHEMA.COLUMNS
WHERE table_name = "customers"
```

1	
2	SELECT
3	column_name,
4	data_type
5	FROM target.INFORMATION_SCHEMA.COLUMNS
6	WHERE table_name = "customers"
7	

Query results			
JOB INFORMATION		RESULTS	EXECUTION DETAILS
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	

**Insight:** Most of the columns inside “customers” table are of String Datatype. It mostly contains personal data about the customer like address.

**B. Get the time range between which the orders were placed.**

**Ans:**

```
SELECT
  MIN(order_purchase_timestamp) AS firstOrder,
  MAX(order_purchase_timestamp) AS lastOrder
FROM `case-study-july-23.target.orders`
```

1	SELECT
2	MIN(order_purchase_timestamp) AS firstOrder,
3	MAX(order_purchase_timestamp) AS lastOrder
4	FROM `case-study-july-23.target.orders`

Query results			
JOB INFORMATION		RESULTS	JSON
		EXECUTION DETAILS	
Row	firstOrder	lastOrder	
1	2016-09-04 21:15:19 UTC	2018-10-17 17:30:18 UTC	

**Insight:** The first order was placed on 04-09-2016 and the last order was placed on 17-10-2018. So, it's been more than 2 years since Target is operating business in the country.

**C. Count the number of Cities and States in our dataset.****Ans:**

```

SELECT
  COUNT(DISTINCT geolocation_state) AS stateCount,
  COUNT(DISTINCT geolocation_city) AS cityCount
FROM `case-study-july-23.target.geolocation`

```

1	SELECT
2	COUNT(DISTINCT geolocation_state) AS stateCount,
3	COUNT(DISTINCT geolocation_city) AS cityCount
4	FROM `case-study-july-23.target.geolocation`
5	

Query results			
JOB INFORMATION		RESULTS	JSON
Row	stateCount	cityCount	EXECUTION DETA
1	27	8011	

**Insight:** Using “Geolocations” table we can conclude that Target is operating business in 28 states and 8011 across Brazil.

## II. In-depth Exploration:

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

Ans:

```
SELECT
  EXTRACT(YEAR FROM order_purchase_timestamp) AS orderYear,
  EXTRACT(MONTH FROM order_purchase_timestamp) AS orderMonth,
  COUNT(order_id) AS totalOrdersPlaced
FROM `case-study-july-23.target.orders`
GROUP BY 1,2
ORDER BY 1,2
```

Row	orderYear	orderMonth	totalOrdersPlaced
1	2016	9	4
2	2016	10	324
3	2016	12	1
4	2017	1	800
5	2017	2	1780
6	2017	3	2682
7	2017	4	2404
8	2017	5	3700
9	2017	6	3245
10	2017	7	4026

**Insight:** Initially the number of orders placed was low but over the years it kept increasing. There are bit fluctuations but still we can conclude that Target has increased sales over the years.

**Recommendation:** The numbers are growing over the years and with better marketing and proper advertising it can grow even more.

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

**Ans:**

```
WITH cte AS
(
  SELECT
    EXTRACT(YEAR FROM order_purchase_timestamp) AS orderYear,
    EXTRACT(MONTH FROM order_purchase_timestamp) AS orderMonth,
    COUNT(order_id) AS totalOrdersPlaced
  FROM `case-study-july-23.target.orders`
  GROUP BY 1,2
)
SELECT
  *,
  DENSE_RANK() OVER(PARTITION BY orderYear ORDER BY totalOrdersPlaced DESC) as rnk
FROM cte
ORDER BY 1
```

Row	orderYear	orderMonth	totalOrdersPlaced	rnk
4	2017	11	7544	1
5	2017	12	5673	2
6	2017	10	4631	3
7	2017	8	4331	4
8	2017	9	4285	5
9	2017	7	4026	6
10	2017	5	3700	7
11	2017	6	3245	8
12	2017	3	2682	9
13	2017	4	2404	10

**Insight 1:** Looking at the monthly sales from year “2017” it’s clear that most of the Brazilian people are shopping more during the last quarter of the year. More precisely we can say that sales maxed during festive period like Christmas and Halloween.

Row	orderYear	orderMonth	totalOrdersPlaced	rnk
16	2018	1	7269	1
17	2018	3	7211	2
18	2018	4	6939	3
19	2018	5	6873	4
20	2018	2	6728	5
21	2018	8	6512	6
22	2018	7	6292	7
23	2018	6	6167	8
24	2018	9	16	9
25	2018	10	4	10

**Insight 2:** As we entered the year “2018” sales were good initially continuing from the festive season in “2017” and then decreased slowly over the months.

So, the best months for Target in Brazil over the years were Nov’17, Jan’18 and Mar’18.

**Recommendation:** Target should focus on the months which have less sales. To improve sales during these months we can offer special discounts. During the months when the sales are high, we can stock more products and similarly we can free up the inventory when sales are less.

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

- 0-6 hrs : Dawn
- 7-12 hrs : Mornings



- 13-18 hrs : Afternoon
- 19-23 hrs : Night

Ans:

```
WITH cte AS
(
  SELECT
    EXTRACT(HOUR FROM order_purchase_timestamp) AS orderHour,
    COUNT(order_id) AS ordersPlaced
  FROM `case-study-july-23.target.orders`
  GROUP BY 1)
SELECT
  CASE
    WHEN orderHour BETWEEN 0 AND 6 THEN "Dawn"
    WHEN orderHour BETWEEN 7 AND 12 THEN "Mornings"
    WHEN orderHour BETWEEN 13 AND 18 THEN "Afternoon"
    WHEN orderHour BETWEEN 19 AND 23 THEN "Night"
  END AS timePeriod,
  SUM(ordersPlaced) as totalOrdersPlaced
FROM cte
GROUP BY timePeriod
ORDER BY 2 DESC
```

Row	timePeriod	totalOrdersPlaced
1	Afternoon	38135
2	Night	28331
3	Mornings	27733
4	Dawn	5242

**Insight:** Data suggests that majority of Brazilian people tend to do shopping in Afternoon. When comes to Night and Mornings numbers are close which means similar amount of people prefer this period.

**Recommendation:** Since the number of orders placed is low during the period Dawn. Target should give discounts and special offers in these periods. This will attract more customers.

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

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

Ans:

```
WITH cte AS
(SELECT
  c.customer_state AS State,
  EXTRACT(MONTH FROM o.order_purchase_timestamp) AS Month,
  COUNT(o.order_id) totalOrdersPlaced
FROM `case-study-july-23.target.orders` o
JOIN `case-study-july-23.target.customers` c
ON o.customer_id = c.customer_id
GROUP BY
  c.customer_state,
  EXTRACT(MONTH FROM o.order_purchase_timestamp))
SELECT *
FROM cte
ORDER BY 1,2
```

Row	State	Month	totalOrdersPlaced
1	AC	1	8
2	AC	2	6
3	AC	3	4
4	AC	4	9
5	AC	5	10
6	AC	6	7
7	AC	7	9
8	AC	8	7
9	AC	9	5
10	AC	10	6

**Insight:** The state of Acre (AC) has the maximum sales in May month followed by April and July. Similarly state Bahia (BA) has the maximum sales in July and May. Using this data, we can conclude that different states have different patterns when it comes to monthly sales. There might be different reasons for that like regional festivals, weather patterns etc.

**Recommendation:** Better inventory management during peak and trough months can optimize the business to a good level. Also, we could focus on product popularity among the states which will also improve inventory.

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

Ans:

```
SELECT
  customer_state as State,
  COUNT(DISTINCT customer_id) as distinctCustomers
FROM `case-study-july-23.target.customers`
GROUP BY customer_state
ORDER BY 2 DESC
```

Row	State	distinctCustomers
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

**Insight 1:** The states like São Paulo, Rio de Janeiro and Minas Gerais have the maximum number of customers. These states have urban populations with better HDI and income per capita.

Row	State	distinctCustomers
18	PI	495
19	RN	485
20	AL	413
21	SE	350
22	TO	280
23	RO	253
24	AM	148
25	AC	81
26	AP	68
27	RR	46

**Insight 2:** The states like Roraima, Amapá and Acre have relatively low customer count since these are not very metropolitan and are mostly inland states. Most of these low customer states are part of Amazon River basin which is a dense tropical rainforest.

**Recommendation:** Target can use targeted marketing for sparsely populated states which have fewer customers. Collaborate with local businesses or organizations in these states to tap into their existing customer networks. These partnerships can help you gain trust and visibility in the region. Consider offering products that align with the environmental consciousness of the Amazon region. Sustainable and eco-friendly products may find a receptive audience in these areas.

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

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

**Ans:**

```
WITH cte AS
(SELECT
  EXTRACT(YEAR FROM order_purchase_timestamp) AS Year,
  ROUND(SUM(payment_value),2) AS totalCostOfOrders
FROM `case-study-july-23.target.orders` o
JOIN `case-study-july-23.target.payments` p
ON o.order_id = p.order_id
WHERE
  EXTRACT(YEAR FROM order_purchase_timestamp) IN (2017,2018)
  AND
  EXTRACT(MONTH FROM order_purchase_timestamp) BETWEEN 1 AND 8
GROUP BY EXTRACT(YEAR FROM order_purchase_timestamp)),
cte2 AS
(SELECT
  Year,
  totalCostOfOrders,
  LAG(totalCostOfOrders) OVER(ORDER BY totalCostOfOrders) as prev_total
FROM cte)
SELECT
  cte2.Year,
  cte2.totalCostOfOrders,
  ROUND((totalCostOfOrders-prev_total)/prev_total*100,2) AS percentageIncrease
FROM cte2
```

Row	Year	totalCostOfOrders	percentageIncrease
1	2017	3669022.12	null
2	2018	8694733.84	136.98

**Insight:** Sales in 2018 were more than double that of 2017 which is nearly an increase of 137%. This establishes that Target has performed quite well in 2018 relatively especially between the given period which is January to August.

**Recommendation:** Optimize inventory levels to ensure products are readily available to meet customer demand without excessive overstocking. Explore opportunities to expand into new markets or demographics where there is potential for growth.

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

**Ans:**

```
SELECT
  customer_state AS state,
  ROUND(SUM(price),2) AS totalOrderPrice,
  ROUND(AVG(price),2) AS avgOrderPrice
FROM `case-study-july-23.target.customers` c
JOIN `case-study-july-23.target.orders` o
ON c.customer_id = o.customer_id
JOIN `case-study-july-23.target.order_items` oi
ON o.order_id = oi.order_id
GROUP BY customer_state
ORDER BY 2 DESC,3 DESC
```

Row	state	totalOrderPrice	avgOrderPrice
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
12	CE	227254.71	153.76
13	PA	178947.81	165.69
14	MT	156453.53	148.3
15	MA	119648.22	145.2

**Insight 1:** The states like São Paulo, Rio de Janeiro and Minas Gerais have the highest total order price since these are metropolitan regions. More population means more business whereas states like Roraima, Amapá and Acre have low total order prices.

**Insight 2:** But when it comes to average order price among the states data tells a different story. The states of Paraíba and Alagoas have the best average order price. This means urban or rural regions do not make much difference when it comes to average price.

**Recommendation:** Given that states like São Paulo, Rio de Janeiro, and Minas Gerais have the highest total order prices due to their large populations and urban nature, consider expanding and intensifying marketing and sales efforts in these regions. While states like Roraima, Amapá, and Acre may have lower total order prices, they still represent opportunities for growth. Explore strategies to penetrate these markets, such as offering competitive pricing, conducting market research to understand local preferences, and improving distribution networks.

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

Ans:

```
SELECT
  customer_state AS state,
  ROUND(SUM(freight_value),2) AS totalFreightValue,
  ROUND(AVG(freight_value),2) AS avgFreightValue
FROM `case-study-july-23.target.customers` c
JOIN `case-study-july-23.target.orders` o
ON c.customer_id = o.customer_id
JOIN `case-study-july-23.target.order_items` oi
ON o.order_id = oi.order_id
GROUP BY customer_state
ORDER BY 2 DESC,3 DESC
```

Row	state	totalFreightValue	avgFreightValue
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

23	RO	11417.38	41.07
24	AM	5478.89	33.21
25	AC	3686.75	40.07
26	AP	2788.5	34.01
27	RR	2235.19	42.98

**Insight 1:** Total freight value is higher in urban states like São Paulo, Rio de Janeiro, and Minas Gerais because of more customers plus more variety of products. Whereas states like Roraima, Amapá and Acre have low total freight value because of less customers.

**Insight 2:** Average freight value in urban states is relatively low because of better connectivity. Plus, most of the warehouses are in metro cities only. But when it comes to the states like Roraima, Amapá and Acre avg freight value is very high due to harsh terrain and poor connectivity.

**Recommendation:** Consider expanding warehouse capacity or establishing additional distribution centers in urban areas to meet the demand. This can reduce lead times and transportation costs. Implement loyalty programs, subscription services, or other customer retention strategies to keep urban customers engaged and encourage repeat purchases.

In states like Roraima, Amapá, and Acre, where total freight value is low due to a smaller customer base and challenging terrain, consider investing in infrastructure improvements to enhance connectivity. Adjust freight pricing in rural areas to reflect the higher costs associated with transportation in these regions. Offer transparent pricing structures to avoid customer dissatisfaction.



## V. Analysis based on sales, freight, and delivery time.

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

Ans:

```
SELECT
  DISTINCT 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,
  CASE
    WHEN
TIMESTAMP_DIFF(order_estimated_delivery_date,order_delivered_customer_date,DAY) < 0
    THEN "Delivered Late"
    WHEN
TIMESTAMP_DIFF(order_estimated_delivery_date,order_delivered_customer_date,DAY) > 0
    THEN "Delivered Early"
    ELSE "Delivered on time"
  END as Delivery_Status
FROM `target.orders`
WHERE order_delivered_customer_date IS NOT NULL
ORDER BY 1
```

Row	order_id	time_to_deliver	diff_estimated_delivery	Delivery_Status
41	001ac194d4a326a6fa99b581e9a3d963	8	4	Delivered Early
42	001b76dd48a5b1eee3e87778daa40df8	11	20	Delivered Early
43	001c85b5f68d2be0cb0797afc9e8ce9a	27	-8	Delivered Late
44	001d8f0e34a38c37f7dba2a37d4eba8b	11	-2	Delivered Late
45	001daeb0eddc45b999bad0801ad9d273	21	10	Delivered Early
46	001dbc16dc51075e987543d23a0507c7	15	34	Delivered Early
47	001e7ba991be1b19605ca0316e7130f9	10	14	Delivered Early
48	001e7cf2ad6bef3ade12ebc56ceaf0f3	16	0	Delivered on time
49	0020262c8a370bd5a174ea6a2a267321	4	19	Delivered Early
50	0020a222f55eb79a372d0efee3cca688	9	12	Delivered Early

**Insight:** Looking at the data we can say that most of the orders are either delivered before the estimated delivery time or on time. But there are some products which are being delivered after estimated delivery time (Negative values represents late delivery).

**Recommendation:** Conduct a thorough analysis to identify which products or categories are most frequently associated with late deliveries. This can help you pinpoint specific areas for improvement. Provide customers with real-time tracking information so they can monitor the progress of their deliveries and receive timely updates.

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

**Ans:**

SELECT

X.state AS topState1,  
X.avgFreightValue AS highestAvgFreightValue,  
Y.state AS topState2,  
Y.avgFreightValue AS lowestAvgFreightValue

FROM

(SELECT  
customer\_state AS state,  
ROUND(AVG(freight\_value),2) AS avgFreightValue,  
DENSE\_RANK() OVER(ORDER BY AVG(freight\_value) DESC) AS ranking  
FROM `target.customers` c  
JOIN `target.orders` o  
ON c.customer\_id = o.customer\_id  
JOIN `target.order\_items` oi  
ON o.order\_id = oi.order\_id  
GROUP BY customer\_state  
ORDER BY 3  
LIMIT 5) AS X

JOIN

(SELECT  
customer\_state AS state,  
ROUND(AVG(freight\_value),2) AS avgFreightValue,  
DENSE\_RANK() OVER(ORDER BY AVG(freight\_value)) AS ranking  
FROM `target.customers` c

```

JOIN `target.orders` o
ON c.customer_id = o.customer_id
JOIN `target.order_items` oi
ON o.order_id = oi.order_id
GROUP BY customer_state
ORDER BY 3
LIMIT 5) AS Y
ON X.ranking = Y.ranking

```

Row	topState1	highestAvgFreightValue	topState2	lowestAvgFreightValue
1	RR	42.98	SP	15.15
2	PB	42.72	PR	20.53
3	RO	41.07	MG	20.63
4	AC	40.07	RJ	20.96
5	PI	39.15	DF	21.04

**Insight:** States like Roraima (RR), Paraíba (PB), Rondônia (RO), Acre (AC) and Piauí (PI) are among the top 5 states with highest average freight value because these are non-urban areas with harsh terrain and bad connectivity.

States like São Paulo (SP), Paraná (PR), Minas Gerais (MG), Rio de Janeiro (RJ) and Distrito Federal (DF) are among the top 5 states with lowest average freight value because these are urban areas with good connectivity and bigger population.

**Recommendation:** In urban states, focus on optimizing delivery routes to minimize congestion and ensure on-time deliveries. Real-time traffic monitoring and adjustment of routes can be valuable. Establish strategically located warehouses in both urban and non-urban areas to minimize transit times and improve efficiency.

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

```

SELECT
  X.state AS topState1,
  X.avgDeliveryTime AS highestAvgDeliveryTime,
  Y.state AS topState2,
  Y.avgDeliveryTime AS lowestAvgDeliveryTime
FROM
  (SELECT
    customer_state AS state,
    ROUND(AVG(TIMESTAMP_DIFF(order_delivered_customer_date,order_purchase_timestamp
,DAY)),2) AS avgDeliveryTime,
    DENSE_RANK() OVER(ORDER BY
AVG(TIMESTAMP_DIFF(order_delivered_customer_date,order_purchase_timestamp,DAY))
DESC) AS
      ranking
    FROM `target.customers` c
    JOIN `target.orders` o
    ON c.customer_id = o.customer_id
    GROUP BY customer_state
    ORDER BY 3
    LIMIT 5) AS X
JOIN
  (SELECT
    customer_state AS state,
    ROUND(AVG(TIMESTAMP_DIFF(order_delivered_customer_date,order_purchase_timestamp
,DAY)),2) AS avgDeliveryTime,
    DENSE_RANK() OVER(ORDER BY
AVG(TIMESTAMP_DIFF(order_delivered_customer_date,order_purchase_timestamp,DAY)))
AS
      ranking
    FROM `target.customers` c
    JOIN `target.orders` o
    ON c.customer_id = o.customer_id
    GROUP BY customer_state
    ORDER BY 3
    LIMIT 5) AS Y
ON X.ranking = Y.ranking

```

Row	topState1	highestAvgDeliveryTime	topState2	lowestAvgDeliveryTime
1	RR	28.98	SP	8.3
2	AP	26.73	PR	11.53
3	AM	25.99	MG	11.54
4	AL	24.04	DF	12.51
5	PA	23.32	SC	14.48

**Insight:** States like Roraima (RR), Amapá (AP), Amazonas (AM), Alagoas (AL) and Pará (PA) have the highest average delivery time due to harsh terrain of Amazon rainforest and poor connectivity. These states are mostly inland or reaching there will take a lot of time.

States like São Paulo (SP), Paraná (PR), Minas Gerais (MG), Distrito Federal (DF) and Santa Catarina (SC) have lowest average delivery time due to better connectivity and these are urban states with better mode of transportations as well.

**Recommendation:** All-terrain vehicles or boats may be necessary in some cases. Implement effective inventory management practices to ensure that products are readily available in these states. In states with poor road connectivity, explore alternative modes of transportation, such as river transport or air freight, to expedite deliveries.

Keep customers informed about expected delivery times and any potential delays. Automated communication systems can provide timely updates.

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

**Ans:**

**SELECT**

```
customer_state AS state,
ROUND(AVG(TIMESTAMP_DIFF(order_estimated_delivery_date,order_delivered_customer_date,DAY)),2) AS avg_actual_vs_expected_DeliveryTime,
DENSE_RANK() OVER(ORDER BY
AVG(TIMESTAMP_DIFF(order_estimated_delivery_date,order_delivered_customer_date,DAY))
DESC) AS ranking
```

```

FROM `target.customers` c
JOIN `target.orders` o
ON c.customer_id = o.customer_id
WHERE order_delivered_customer_date IS NOT NULL
GROUP BY customer_state
ORDER BY 3
LIMIT 5

```

Row	state	avg_actual_vs_expected_DeliveryTime	ranking
1	AC	19.76	1
2	RO	19.13	2
3	AP	18.73	3
4	AM	18.61	4
5	RR	16.41	5

**Insight:** States like Acre (AC), Rondônia (RO), Amapá (AP), Amazonas (AM) and Roraima (RR) have the highest average difference between actual and expected delivery date. The reason this is happening is because the average delivery time for these states is relatively high. So, the estimated delivery time is deliberately kept high for these Inland states and most of the time orders are delivered before estimated time.

**Recommendation:** Utilize historical delivery data to calculate more accurate estimated delivery times for each state. Instead of deliberately overestimating, aim to provide customers with realistic delivery windows.

Encourage customers to provide feedback on their delivery experiences. Use this feedback to make necessary improvements and address any recurring issues.

## VI. Analysis based on the payments:

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

Ans:

```
SELECT
  EXTRACT(MONTH FROM order_purchase_timestamp) AS month,
  payment_type AS paymentMethod,
  COUNT(o.order_id) as totalPurchases
FROM `target.orders` o
JOIN `target.payments` p
ON o.order_id = p.order_id
GROUP BY 1,2
ORDER BY 1
```

Row	month	paymentMethod	totalPurchases
1	1	credit_card	6103
2	1	UPI	1715
3	1	voucher	477
4	1	debit_card	118
5	2	UPI	1723
6	2	credit_card	6609
7	2	voucher	424
8	2	debit_card	82
9	3	credit_card	7707
10	3	UPI	1942
11	3	debit_card	109
12	3	voucher	591
13	4	voucher	572
14	4	credit_card	7301
15	4	UPI	1783

**Insight:** Majority of people are either using credit card or UPI. Most of the people like to go cashless now. Debit cards and Vouchers are also being used but the frequency is relatively low. People do tend to use credit cards because of the exciting benefits it offers plus it also increases their credit score.

**Recommendation:** Offer incentives like discounts, cashback rewards, or loyalty points for customers who choose cashless payment methods. These incentives can motivate individuals to adopt digital payment options.

Ensure that your mobile app or online platform is user-friendly and supports various digital payment methods like UPI, credit cards, and digital wallets. Make the payment process seamless and secure.

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

**Ans:**

```
SELECT
    payment_installments AS noOfInstallments,
    COUNT(o.order_id) AS totalPurchases
FROM `target.orders` o
JOIN `target.payments` p
ON o.order_id = p.order_id
GROUP BY payment_installments
ORDER BY 2 DESC
```



Row	noOfInstallments	totalPurchases
1	1	52546
2	2	12413
3	3	10461
4	4	7098
5	10	5328
6	5	5239
7	8	4268
8	6	3920
9	7	1626
10	9	644

**Insight:** Based on the data analysis, it becomes evident that the majority of customers opt for single-installment payments. Conversely, individuals who choose the EMI (Equated Monthly Installments) option tend to favor splitting their payments into 2-3 installments. As the number of installments increases, the frequency of purchases decreases. However, notably, the number "10" stands out as a preferred option for installment payments, likely due to its simplicity for calculation purposes.

**Recommendation:** Clearly communicate the terms and conditions associated with each installment plan. Ensure that customers have a complete understanding of any fees, interest rates, and payment schedules.

Ensure that your payment platform simplifies calculations for installment payments, especially for popular choices like "10."

Educate customers about the potential positive impact on their credit scores when they use installment plans responsibly and make payments on time.

**Key takeaways from the case study: -**

1. **Operational Duration:** Target has been operating in Brazil for over two years, with its first order recorded in September 2016 and continuous operations until October 2018.
2. **Geographical Presence:** Target operates in 28 states and over 8,000 cities across Brazil, highlighting the extent of its geographical reach.
3. **Sales Growth:** Over the years, Target has experienced sales growth in Brazil, with minor fluctuations but a consistent upward trend. Marketing and advertising efforts can further boost sales.
4. **Seasonal Sales:** Brazilian customers tend to shop more during the last quarter of the year, particularly during festive periods like Christmas and Halloween.
5. **Preferred Shopping Times:** Afternoon is the most popular time for Brazilian customers to place orders, with mornings and nights being secondary preferences.
6. **State-Specific Sales Patterns:** Different states exhibit varying monthly sales patterns, likely influenced by regional festivals, weather conditions, and cultural factors. Target can tailor its marketing and inventory strategies accordingly.
7. **Payment Preferences:** Brazilian customers primarily prefer digital payment methods like credit cards and UPI.
8. **Inventory Management:** Optimizing inventory levels is crucial to meet customer demand efficiently, especially during peak sales months.
9. **Customer Feedback Loop:** Establish a feedback mechanism to gather customer insights and continuously improve the customer experience, especially in regions with specific challenges.
10. **Economic Impact:** Sales in 2018 saw a significant increase compared to 2017, highlighting Target's strong performance.