**CASE STUDY**

**Target Dataset:**

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.Initial exploration like checking the structure & characteristics of the data:**

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

SELECT COLUMN\_NAME, DATA\_TYPE

FROM data-44-7685.casestudy\_target.INFORMATION\_SCHEMA.COLUMNS

WHERE TABLE\_NAME = 'customers'

A screenshot of a data

AI-generated content may be incorrect.

**Insights:**

* The table contains **customer\_id, customer\_unique\_id, zip code, city, and state**, which helps track customer location and uniqueness.
* customer\_id and customer\_unique\_id are **both STRING** types, indicating that customer records might be referenced by different identifiers.
* customer\_zip\_code\_prefix is stored as INT64, meaning only the prefix of the zip code is captured, **not the full zip code**.

**2. Check for Missing Values**

SELECT

COUNT(\*) - COUNT(customer\_city) AS missing\_values,

customer\_city

FROM `data-44-7685.casestudy\_target.customers`

GROUP BY customer\_city;

A screenshot of a computer

AI-generated content may be incorrect.

There are no missing values in the columns table.

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

SELECT

MIN(order\_purchase\_timestamp) AS first\_order\_time,

MAX(order\_purchase\_timestamp) AS last\_order\_time

FROM `casestudy\_target.orders`;

A screenshot of a computer

AI-generated content may be incorrect.

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

SELECT

COUNT(DISTINCT customer\_city) AS unique\_cities,

COUNT(DISTINCT customer\_state) AS unique\_states

FROM `casestudy\_target.customers` c inner join

`casestudy\_target.orders` o on c.customer\_id = o.customer\_id

**A screenshot of a computer

AI-generated content may be incorrect.**

Orders have been placed by customers from 4119 unique cities and 27 unique states**.**

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

select

extract(year from order\_purchase\_timestamp) as year,

count(1) as num\_orders

from `casestudy\_target.orders`

group by year

order by year;

A screenshot of a report

AI-generated content may be incorrect.

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

select

extract(year from order\_purchase\_timestamp) as year,

extract(month from order\_purchase\_timestamp) as month,

count(1) as num\_orders

from `casestudy\_target.orders`

group by year, month

order by year, month;

**A screenshot of a data report

AI-generated content may be incorrect.**

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

* 1. **0-6 hrs : Dawn**
  2. **7-12 hrs : Mornings**
  3. **13-18 hrs : Afternoon**
  4. **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 '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(distinct order\_id) as num\_of\_orders

from `casestudy\_target.orders`

group by time\_of\_day

order by num\_of\_orders desc

A screenshot of a report

AI-generated content may be incorrect.

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

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

SELECT c.customer\_state,

extract(year from order\_purchase\_timestamp) as year,

extract(month from order\_purchase\_timestamp) as month,

count(1) as num\_orders

FROM `casestudy\_target.customers` c inner join

`casestudy\_target.orders` o on c.customer\_id = o.customer\_id

group by 1,2,3

order by 4 desc

**A screenshot of a data sheet

AI-generated content may be incorrect.**

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

SELECT customer\_state,

count(customer\_id) as num\_customers

from `casestudy\_target.customers`

group by 1

order by 2 desc

**A screenshot of a data report

AI-generated content may be incorrect.**

**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.**

with base\_1 as

(

select \* from

`casestudy\_target.orders` o inner join `casestudy\_target.payments` p on o.order\_id = p.order\_id

where extract(year from order\_purchase\_timestamp) between 2017 and 2018 and

extract(month from order\_purchase\_timestamp) between 1 and 8

),

base\_2 as

(

select

extract(year from order\_purchase\_timestamp) as year,

sum(payment\_value) as cost

from base\_1

group by 1

order by 1),

base\_3 as

(

select \*,

lead(cost) over(order by year) as next\_year

from base\_2

)

select \*,

round((next\_year - cost)/cost \* 100,2) as percentage\_increase

from base\_3;

**A screenshot of a computer

AI-generated content may be incorrect.**

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

select c.customer\_state,

sum(price) as total\_price,

avg(price) as avg\_price

from `casestudy\_target.order\_items` oi join `casestudy\_target.orders` o on oi.order\_id = o.order\_id

join `casestudy\_target.customers` c on o.customer\_id = c.customer\_id

group by customer\_state

order by 2 desc

A screenshot of a data

AI-generated content may be incorrect.

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

select c.customer\_state,

round(sum(freight\_value),2) as total\_price,

round(avg(freight\_value),2) as avg\_price

from `casestudy\_target.order\_items` oi join `casestudy\_target.orders` o on oi.order\_id = o.order\_id

join `casestudy\_target.customers` c on o.customer\_id = c.customer\_id

group by customer\_state

order by 2 desc

A screenshot of a data

AI-generated content may be incorrect.

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

**5.1.Find the no. of days taken to deliver each order from the order’s purchase date as delivery time.  
Also, calculate the difference (in days) between the estimated & actual delivery date of an order.**

select

timestamp\_diff(order\_delivered\_customer\_date, order\_purchase\_timestamp, DAY) as time\_to\_deliver,

timestamp\_diff(order\_delivered\_customer\_date, order\_estimated\_delivery\_date, DAY) as diff\_estimated\_delivery

from `casestudy\_target.orders`

where order\_status = 'delivered'

**A screenshot of a data

AI-generated content may be incorrect.**

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

select

c.customer\_state,

round(avg(freight\_value),2) as avg\_freight\_value,

from `casestudy\_target.order\_items` oi join `casestudy\_target.orders` o using(order\_id)

join `casestudy\_target.customers` c using(customer\_id)

group by customer\_state

order by 2 desc

limit 5

**A screenshot of a data

AI-generated content may be incorrect.**

select

c.customer\_state,

round(avg(freight\_value),2) as avg\_freight\_value,

from `casestudy\_target.order\_items` oi join `casestudy\_target.orders` o using(order\_id)

join `casestudy\_target.customers` c using(customer\_id)

group by customer\_state

order by avg\_freight\_value asc

limit 5

**A screenshot of a computer

AI-generated content may be incorrect.**

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

SELECT c.customer\_state,

AVG(date\_diff(order\_delivered\_customer\_date, order\_purchase\_timestamp, DAY)) AS avg\_delivery\_time

FROM `casestudy\_target.orders` o join `casestudy\_target.customers` c using(customer\_id)

GROUP BY customer\_state

ORDER BY avg\_delivery\_time DESC

LIMIT 5;

**A screenshot of a data

AI-generated content may be incorrect.**

SELECT c.customer\_state,

AVG(date\_diff(order\_delivered\_customer\_date, order\_purchase\_timestamp, DAY)) AS avg\_delivery\_time

FROM `casestudy\_target.orders` o join `casestudy\_target.customers` c using(customer\_id)

GROUP BY customer\_state

ORDER BY avg\_delivery\_time asc

LIMIT 5;

**A screenshot of a data

AI-generated content may be incorrect.**

* 1. **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.**

select c.customer\_state,

round(sum(timestamp\_diff(order\_delivered\_customer\_date, order\_purchase\_timestamp, DAY))/ count(order\_id), 2) as avg\_time\_to\_deliver,

round(sum(timestamp\_diff(order\_estimated\_delivery\_date, order\_purchase\_timestamp, DAY)) / count(order\_id), 2) as avg\_diff\_estimated\_delivery

from `casestudy\_target.orders` o join `casestudy\_target.customers` c using(customer\_id)

group by customer\_state

order by (avg\_time\_to\_deliver - avg\_diff\_estimated\_delivery)

**A screenshot of a spreadsheet

AI-generated content may be incorrect.**

1. **Analysis based on the payments:**
2. **Find the month on month no. of orders placed using different payment types.**

select

extract(year from order\_purchase\_timestamp) as year,

extract(month from order\_purchase\_timestamp) as month,

payment\_type,

count(order\_id) as num\_orders

from `casestudy\_target.orders` o join `casestudy\_target.payments` p using(order\_id)

GROUP BY year,month, payment\_type

ORDER BY year,month, payment\_type;

A screenshot of a data

AI-generated content may be incorrect.

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

select payment\_installments as installments,

count(order\_id) as num\_orders

from `casestudy\_target.payments`

where payment\_installments >= 1

group by payment\_installments

order by 2 desc

**A screenshot of a computer

AI-generated content may be incorrect.**

**Insights:**

1. **Customer Distribution and Order Patterns:**

• Orders were placed by customers from 4,119 unique cities across 27 states.

• The order trend shows fluctuations over the years, with certain months showing higher order volumes, indicating seasonality.

• Customers mostly placed orders in the afternoon period, suggesting peak engagement during this time.

2. **Economic Impact and Pricing Trends:**

• There was a significant increase in order costs from 2017 to 2018 (Jan–Aug), highlighting a potential increase in product prices or customer spending.

• States with the highest average order prices contributed more to overall revenue, while some states consistently showed lower spending.

3. **Freight and Delivery Analysis:**

• Certain states had higher average freight values, suggesting logistical challenges or longer distances.

• Delivery times varied by state, with some consistently performing better than others.

• Some states showed faster deliveries ahead of the estimated dates, highlighting efficient logistics in those areas.

4. **Payment Trends:**

• There were variations in payment types used across months, reflecting customer preferences and possible external factors (like promotions or discounts).

• A higher number of orders were made with installment payments, indicating that flexible payment options were popular among customers.

**Recommendations:**

**1.Optimize Logistics and Freight Costs:**

• Focus on states with higher freight costs to optimize shipping methods and reduce expenses.

• Collaborate with local logistics partners for faster, cost-effective delivery.

**2. Enhance Payment Flexibility:**

• Promote installment-based payment options further, especially during peak seasons.

• Offer incentives for faster payment methods to reduce payment delays.

**3. Focus on Peak Ordering Times:**

• Boost afternoon marketing campaigns and promotions, as most orders are placed during this period.

• Align staffing and support services to these peak hours for better customer service.

**4. Improve Delivery Efficiency:**

• Identify and replicate the practices of states with faster deliveries across other regions.

• Set realistic estimated delivery dates to better manage customer expectations.

**5. Monitor Pricing Trends:**

• Evaluate states with consistent price increases to understand the cause (like inflation, supply chain issues, etc.).

• Offer discounts or loyalty benefits in regions where customers show lower spending patterns.