**Business Case: Target SQL**

Table of Contents

[Objective 2](#_Toc113223994)

[Inputs 2](#_Toc113223995)

[Dataset 2](#_Toc113223996)

[Relationship between datasets 2](#_Toc113223997)

[Data Pre-processing 2](#_Toc113223998)

[BigQuery workbench 2](#_Toc113223999)

[Metadata of dataset 2](#_Toc113224000)

[Describe the dataset by Table 3](#_Toc113224001)

[Ordinal position 3](#_Toc113224002)

[Data Exploration & Analysis 4](#_Toc113224003)

[Customers Table 4](#_Toc113224004)

[Sellers Table 5](#_Toc113224005)

[Order Items Table 10](#_Toc113224006)

[Payments Table 17](#_Toc113224007)

[Orders Table 19](#_Toc113224008)

[Reviews Table 23](#_Toc113224009)

[Products Table 24](#_Toc113224010)

[Insights 28](#_Toc113224011)

[Recommendations 29](#_Toc113224012)

# 

# **Objective**

Analyze the given dataset from Target Store, derive insights and present the recommendations.

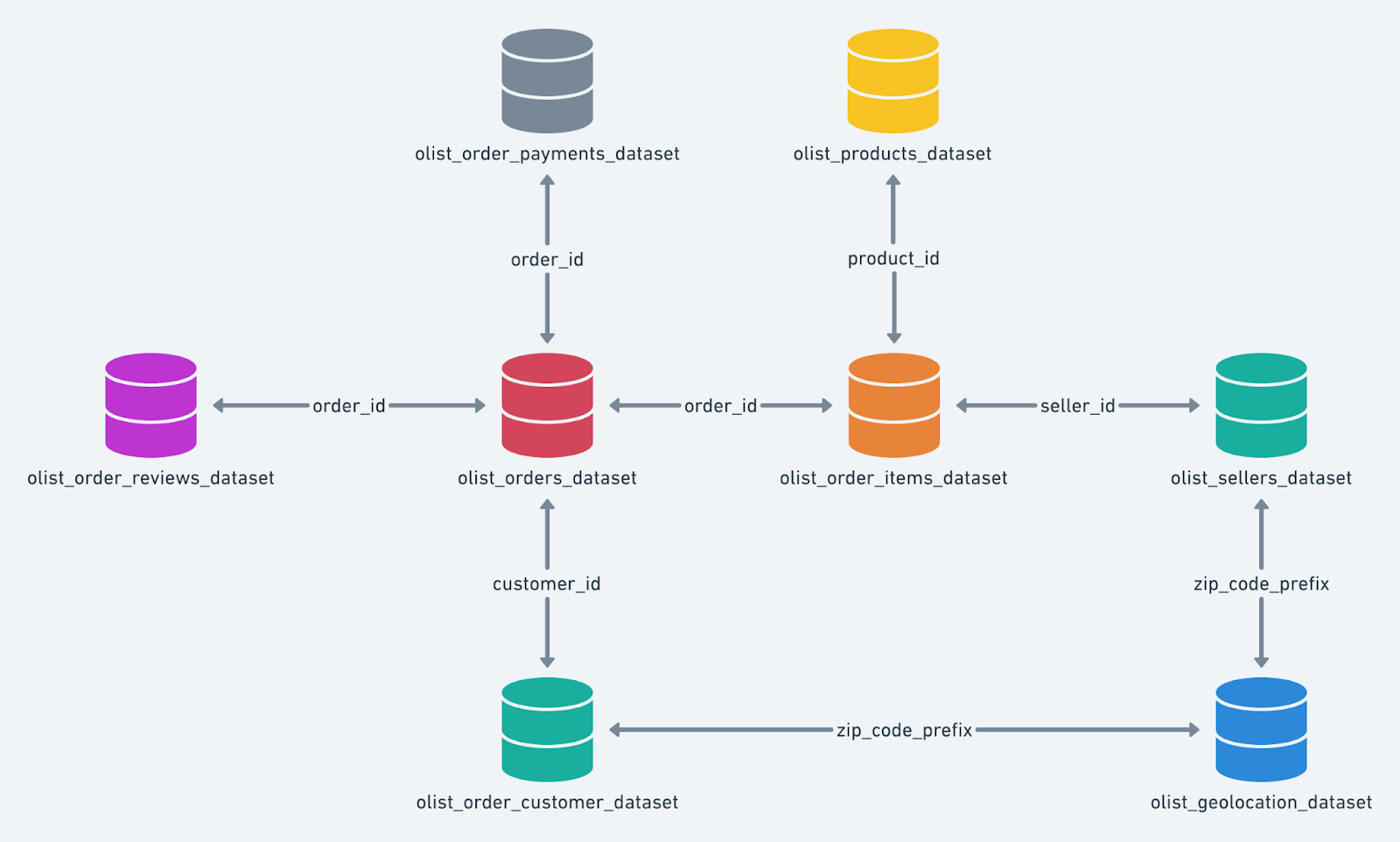
# **Inputs**

This business case has information of 100k orders from 2016 to 2018 made at Target in Brazil. Data is available in 8 csv files about customers, geolocation, order\_items, payments, reviews, orders, products, and sellers.

## **Dataset**

<https://drive.google.com/drive/folders/1TGEc66YKbD443nslRi1bWgVd238gJCnb>

## **Relationship between datasets**



# **Data Pre-processing**

## **BigQuery workbench**

<https://console.cloud.google.com/bigquery?project=target-sql-361009>

## **Metadata of dataset**

SELECT \* EXCEPT(is\_generated, generation\_expression, is\_stored, is\_updatable)

FROM `target-sql-361009.TargetDS.INFORMATION\_SCHEMA.COLUMNS`

WHERE table\_name = ‘customers’

SELECT \* EXCEPT(is\_generated, generation\_expression, is\_stored, is\_updatable)

FROM `target-sql-361009.TargetDS.INFORMATION\_SCHEMA.COLUMNS`

WHERE table\_name = 'orders'

SELECT \* EXCEPT(is\_generated, generation\_expression, is\_stored, is\_updatable)

FROM `target-sql-361009.TargetDS.INFORMATION\_SCHEMA.COLUMNS`

WHERE table\_name = 'payments'

SELECT \* EXCEPT(is\_generated, generation\_expression, is\_stored, is\_updatable)

FROM `target-sql-361009.TargetDS.INFORMATION\_SCHEMA.COLUMNS`

WHERE table\_name = 'products'

## **Describe the dataset by Table**

DECLARE columns ARRAY<STRING>;

SET columns = (

WITH all\_columns AS (

SELECT column\_name

FROM `target-sql-361009.TargetDS.INFORMATION\_SCHEMA.COLUMNS`

WHERE table\_name = 'products'

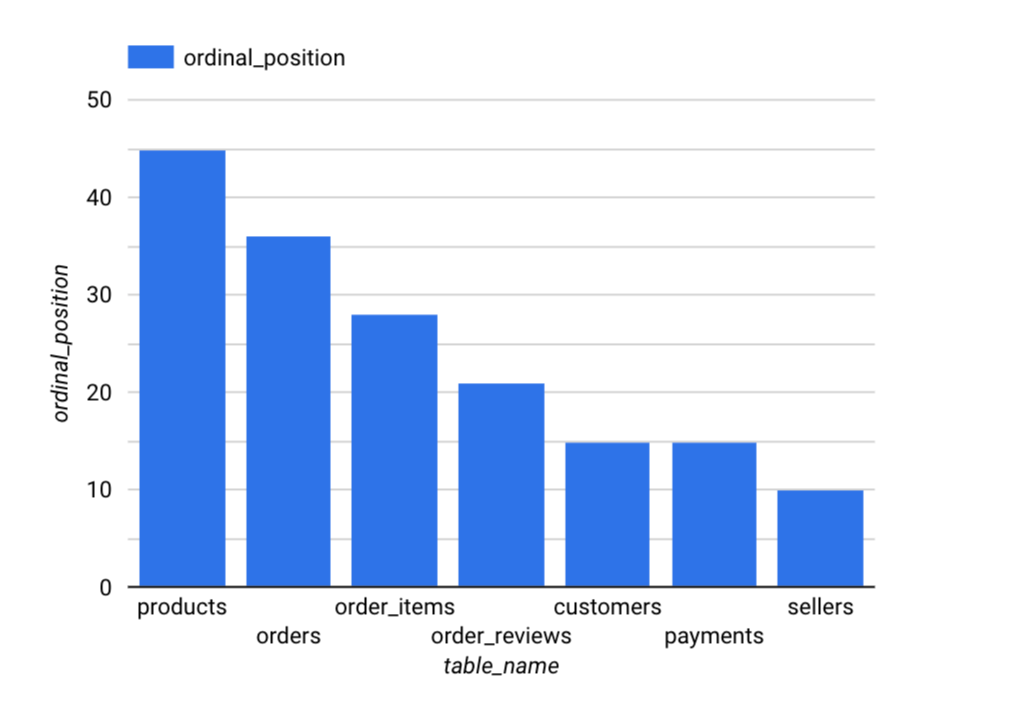
and data\_type IN ('INT64','FLOAT64'))

SELECT ARRAY\_AGG((column\_name) ) AS columns FROM all\_columns );

## **Ordinal position**

SELECT \* FROM `target-sql-361009.TargetDS.INFORMATION\_SCHEMA.COLUMNS`





Data from geo\_locations table may not be required as the same data available from other given tables.

# **Data Exploration & Analysis**

## **Customers Table**

1. Total Customers count

SELECT COUNT(customer\_unique\_id)

FROM `target-sql-361009.TargetDS.customers`

**99441**

1. Total Cities count

SELECT count(DISTINCT(customer\_city))

FROM `target-sql-361009.TargetDS.customers`

**4119**

1. Total States count

SELECT count(DISTINCT(customer\_state))

FROM `target-sql-361009.TargetDS.customers`

**27**

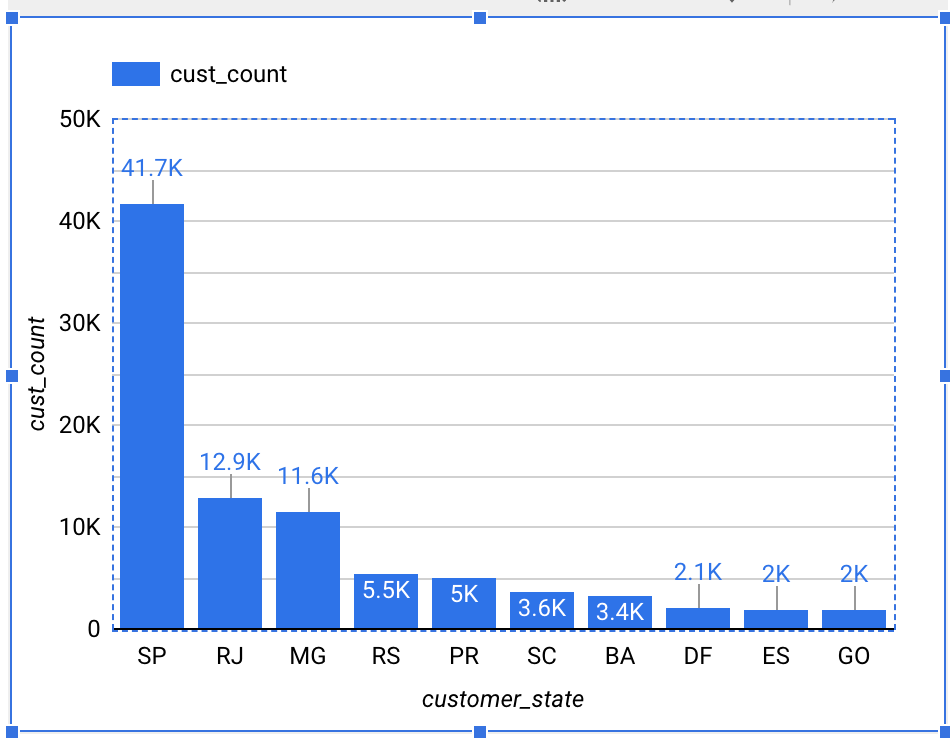
1. Top 10 States by Customers count

SELECT customer\_state, COUNT(customer\_id) AS cust\_count

FROM `target-sql-361009.TargetDS.customers`

GROUP BY customer\_state

ORDER BY cust\_count DESC LIMIT 10



The states of San Paulo and Rio de janeiro have more Customers

1. Top 10 Cities in terms of Customer Count

SELECT customer\_city, COUNT(customer\_unique\_id) cust\_count

FROM `target-sql-361009.TargetDS.customers`

GROUP BY customer\_city

ORDER BY cust\_count DESC LIMIT 10

A screen shot of a graph

Description automatically generatedIn line with the States, the respective cities indicate they have the high number of customers.

## **Sellers Table**

1. Total Sellers count

SELECT count(seller\_id)

FROM `target-sql-361009.TargetDS.sellers`

**3095**

1. Total Cities count

SELECT COUNT(DISTINCT(seller\_city))

FROM `target-sql-361009.TargetDS.sellers`

**611**

1. Total States count

SELECT COUNT(DISTINCT(seller\_state))

FROM `target-sql-361009.TargetDS.sellers`

**23**

1. Top 10 Cities with high number of Sellers

SELECT seller\_city, count(seller\_id) AS seller\_cnt

FROM `target-sql-361009.TargetDS.sellers`

GROUP BY seller\_city

ORDER BY seller\_cnt DESC LIMIT 10

A screenshot of a computer

Description automatically generated

1. Top 10 States with high number of Sellers

SELECT seller\_state, count(seller\_id) AS seller\_cnt

FROM `target-sql-361009.TargetDS.sellers`

GROUP BY seller\_state

ORDER BY seller\_cnt DESC LIMIT 10

A screenshot of a computer

Description automatically generatedStates of San Paulo, Curitiba and Rio continue to be on top even in terms of Sellers count.

1. Top 10 Cities of Sellers with high number of Orders

SELECT seller\_city, count(order\_item\_id) order\_cnt

FROM `target-sql-361009.TargetDS.sellers` sel

JOIN `target-sql-361009.TargetDS.order\_items` order\_item

ON sel.seller\_id = order\_item.seller\_id

GROUP BY seller\_city

ORDER BY order\_cnt DESC LIMIT 10

A screen shot of a graph

Description automatically generated

1. Top 10 Cities of Sellers with high number of Orders

SELECT seller\_city, count(distinct(sel.seller\_id)) sel\_cnt, count(order\_item\_id) order\_cnt

FROM `target-sql-361009.TargetDS.sellers` sel

JOIN `target-sql-361009.TargetDS.order\_items` order\_item

ON sel.seller\_id = order\_item.seller\_id

GROUP BY seller\_city

ORDER BY order\_cnt DESC LIMIT 10

A screenshot of a results table

Description automatically generated

Though Rio has more number of Sellers, but recorded less number of Orders as compared to other cities.

1. Top 10 Cities with low number of Sellers their respective count of Orders

SELECT seller\_city, count(distinct(sel.seller\_id)) sel\_cnt, count(order\_item\_id) order\_cnt

FROM `target-sql-361009.TargetDS.sellers` sel

ON sel.seller\_id = order\_item.seller\_id

GROUP BY seller\_city

ORDER BY order\_cnt ASC LIMIT 10

A screenshot of a computer

Description automatically generated

The cities with low count of Sellers have also resulted in low number of Orders

1. Top 10 States of Sellers with high number of Orders

SELECT seller\_state, count(distinct(sel.seller\_id)) sel\_cnt, count(order\_item\_id) order\_cnt

FROM `target-sql-361009.TargetDS.sellers` sel

JOIN `target-sql-361009.TargetDS.order\_items` order\_item

ON sel.seller\_id = order\_item.seller\_id

GROUP BY seller\_state

ORDER BY order\_cnt DESC LIMIT 10

A screenshot of a data report

Description automatically generated

High number of Sellers also have resulted in high number of Orders in the states.

1. States of less number of Sellers and their respective number of Orders

SELECT seller\_state, count(distinct(sel.seller\_id)) sel\_cnt, count(order\_item\_id) order\_cnt

FROM `target-sql-361009.TargetDS.sellers` sel

JOIN `target-sql-361009.TargetDS.order\_items` order\_item

ON sel.seller\_id = order\_item.seller\_id

GROUP BY seller\_state

ORDER BY order\_cnt ASC LIMIT 10

A screenshot of a computer

Description automatically generated

The states with less number of Sellers have recorded low orders.

1. Top 10 States of Sellers with high number of Customers

SELECT seller\_state, count(distinct(sel.seller\_id)) sel\_cnt, count(distinct(customer\_id)) cust\_cnt

FROM `target-sql-361009.TargetDS.sellers` sel

JOIN `target-sql-361009.TargetDS.order\_items` order\_item

ON sel.seller\_id = order\_item.seller\_id

JOIN `target-sql-361009.TargetDS.orders` ord

ON order\_item.order\_id=ord.order\_id

GROUP BY seller\_state

ORDER BY cust\_cnt DESC LIMIT 10

A screenshot of a data report

Description automatically generated

The states with high number of Sellers also seem to have high number of Customers.

1. States of less number of Sellers and their respective number of Orders

SELECT seller\_state, count(distinct(sel.seller\_id)) sel\_cnt, count(distinct(customer\_id)) cust\_cnt

FROM `target-sql-361009.TargetDS.sellers` sel

JOIN `target-sql-361009.TargetDS.order\_items` order\_item

ON sel.seller\_id = order\_item.seller\_id

JOIN `target-sql-361009.TargetDS.orders` ord

ON order\_item.order\_id=ord.order\_id

GROUP BY seller\_state

ORDER BY cust\_cnt ASC LIMIT 10

A screenshot of a computer

Description automatically generated

The state of ‘AC’ seem to have only 1 Seller and 1 Customer.

## **Order Items Table**

1. Count of Orders, Order Items, Products and Sellers

SELECT count(DISTINCT(order\_id)) AS OrderCnt, count(DISTINCT(order\_item\_id)) AS OrderItemCnt, count(DISTINCT(product\_id)) AS ProductCnt, count(DISTINCT(seller\_id)) AS SellerCnt

FROM `target-sql-361009.TargetDS.order\_items`

A screenshot of a computer

Description automatically generated

1. Top 10 Products in terms of Price

SELECT product\_id, price

FROM `target-sql-361009.TargetDS.order\_items`

ORDER BY price DESC LIMIT 10

A screenshot of a data report

Description automatically generated

1. Top 10 Sellers

SELECT seller\_id, count(distinct(order\_id)) AS OrderId

FROM `target-sql-361009.TargetDS.order\_items`

GROUP BY seller\_id

ORDER BY OrderId DESC LIMIT 10

A screenshot of a computer

Description automatically generated

1. Products that are sold more

SELECT product\_id, count(order\_id) AS OrderId

FROM `target-sql-361009.TargetDS.order\_items`

GROUP BY product\_id

ORDER BY OrderId DESC LIMIT 10

A screenshot of a data report

Description automatically generated

1. Are there any Sellers who didn’t have any sales? - No

SELECT seller\_id

FROM `target-sql-361009.TargetDS.order\_items`

WHERE order\_id is null

The above query returned zero results, which means all the registered sellers at Target are active and they have sales.

1. Seller who has maximum number of orders

SELECT seller\_id, count(order\_id) AS order\_cnt

FROM `target-sql-361009.TargetDS.order\_items`

GROUP BY seller\_id

ORDER BY order\_cnt DESC LIMIT 1

A screenshot of a computer

Description automatically generated

1. Seller with minimum number of orders

SELECT seller\_id, count(order\_id) AS order\_cnt

FROM `target-sql-361009.TargetDS.order\_items`

GROUP BY seller\_id

ORDER BY order\_cnt ASC LIMIT 1

A screenshot of a computer

Description automatically generated

1. Seller with maximum freight value

SELECT seller\_id, freight\_value

FROM `target-sql-361009.TargetDS.order\_items`

GROUP BY seller\_id, freight\_value

ORDER BY freight\_value DESC LIMIT 10

A screenshot of a data

Description automatically generated

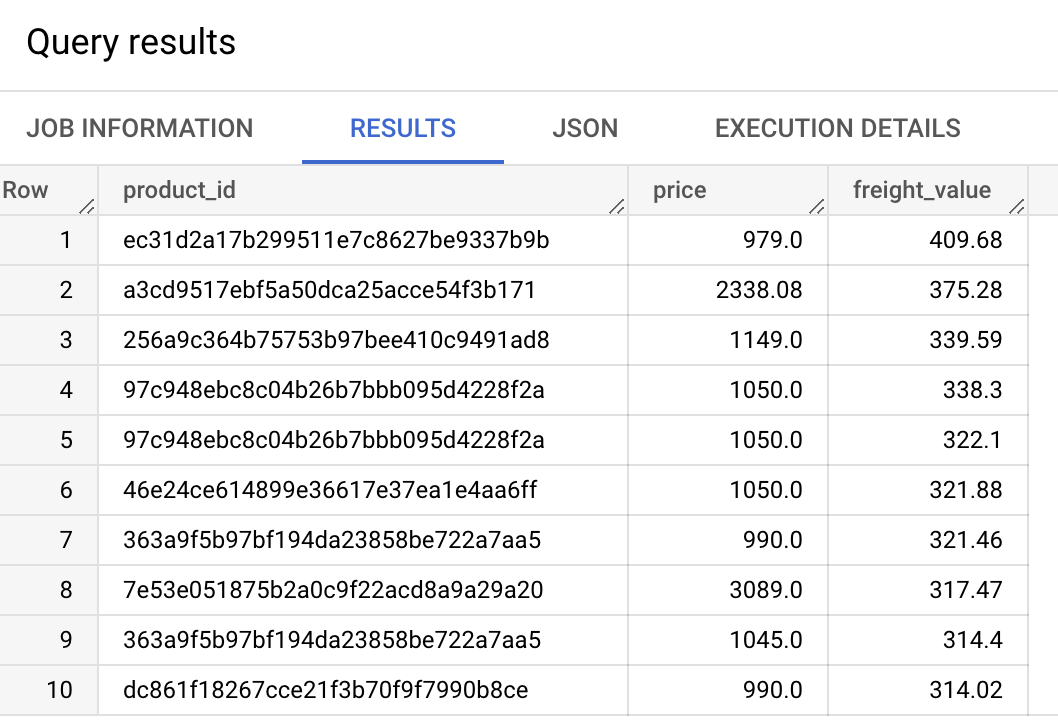
1. Correlation b/w product price and freight value

SELECT product\_id, price, freight\_value

FROM `target-sql-361009.TargetDS.order\_items`

GROUP BY product\_id, price, freight\_value

ORDER BY freight\_value DESC LIMIT 10



There is no correlation b/w the product price and high freight value

1. Seller who charges minimum freight value

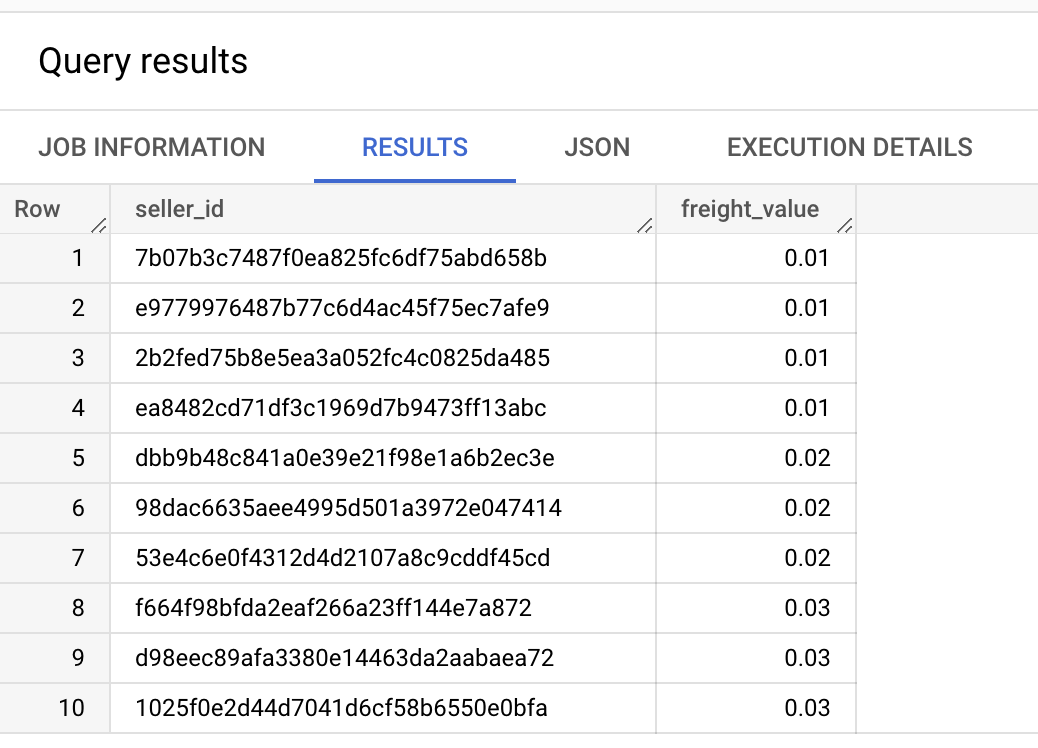
SELECT seller\_id, freight\_value

FROM `target-sql-361009.TargetDS.order\_items`

WHERE freight\_value>0

GROUP BY seller\_id, freight\_value

ORDER BY freight\_value ASC LIMIT 10



1. Sellers with zero freight value

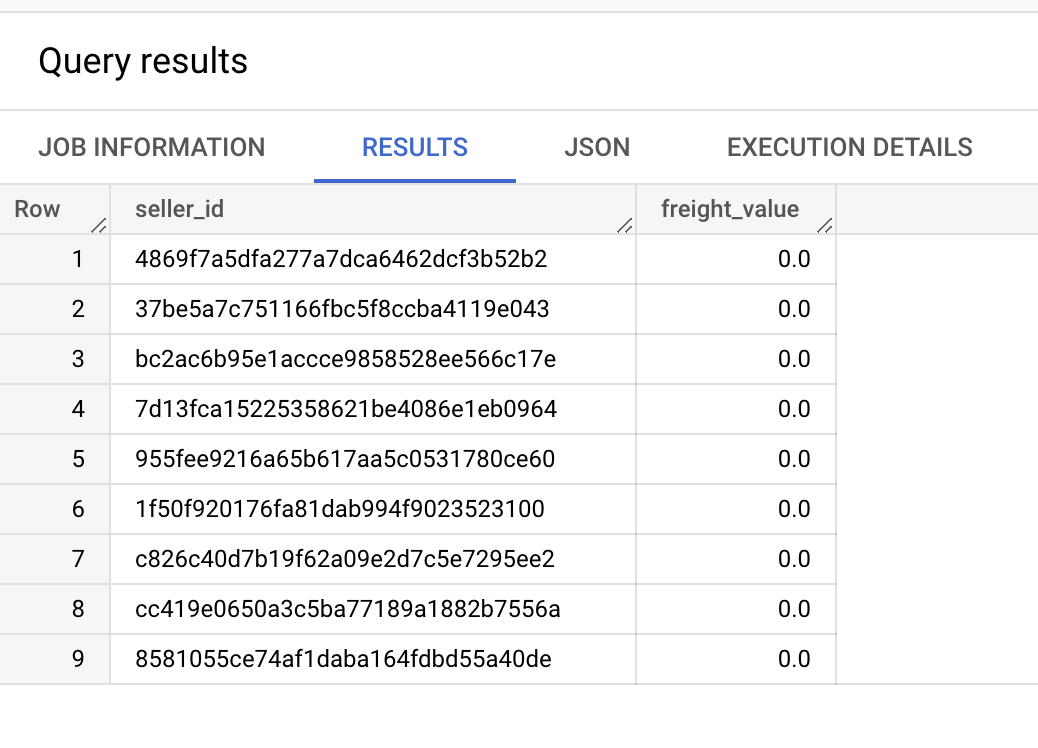
SELECT seller\_id, freight\_value

FROM `target-sql-361009.TargetDS.order\_items`

WHERE freight\_value=0

GROUP BY seller\_id, freight\_value

ORDER BY freight\_value



1. Sellers with zero freight value and their order count

SELECT seller\_id, freight\_value, count(order\_id) AS order\_cnt

FROM `target-sql-361009.TargetDS.order\_items`

WHERE freight\_value=0

GROUP BY seller\_id, freight\_value

ORDER BY order\_cnt DESC

A screenshot of a results report

Description automatically generated

1. Sellers with zero freight value and their order count

SELECT count(distinct(seller\_id)) AS seller\_cnt, count(order\_id) AS order\_cnt

FROM `target-sql-361009.TargetDS.order\_items`

WHERE freight\_value=0

A screenshot of a computer

Description automatically generatedThere are 9 sellers who have zero freight charges, their respective order counts.

1. Sellers with high freight value and their order count

SELECT seller\_id, freight\_value, count(order\_id) AS order\_cnt

FROM `target-sql-361009.TargetDS.order\_items`

GROUP BY seller\_id, freight\_value

ORDER BY freight\_value DESC LIMIT 10

A screenshot of a computer

Description automatically generatedThe sellers with high freight value have less number of orders

1. Number of orders by Sellers whose freight value is above 20

SELECT count(distinct(seller\_id)) AS seller\_cnt, count(distinct(order\_id)) AS order\_cnt

FROM `target-sql-361009.TargetDS.order\_items`

WHERE freight\_value>20

A screenshot of a computer

Description automatically generated30% of orders have freight value $20 and above.

1. Sellers with high number of orders and their freight value

SELECT seller\_id, freight\_value, count(order\_id) AS order\_cnt

FROM `target-sql-361009.TargetDS.order\_items`

GROUP BY seller\_id, freight\_value

ORDER BY order\_cnt DESC

A screenshot of a data report

Description automatically generatedThis indicates that the sellers with relative lower freight value have higher number of orders

1. # of products whose freight value is between 10 and 20 and their order counts

SELECT count(distinct(product\_id)) AS product\_cnt, count(distinct(order\_id)) AS order\_cnt

FROM `target-sql-361009.TargetDS.order\_items`

WHERE freight\_value BETWEEN 10 AND 20

A screenshot of a computer

Description automatically generated

This indicates that 80% of orders are from products who have freight value b/w 10 and 20

1. # of products whose freight value is between 1 and 10 and their order counts

SELECT count(distinct(product\_id)) AS product\_cnt, count(distinct(order\_id)) AS order\_cnt

FROM `target-sql-361009.TargetDS.order\_items`

WHERE freight\_value>0 AND freight\_value<=10

A screenshot of a computer

Description automatically generatedThis indicates that 13% of orders are from products who have freight value less than 10.

## **Payments Table**

1. Total count of orders from Payments table

SELECT count(order\_id) AS total\_order, count(distinct(order\_id)) AS distinct\_order FROM `target-sql-361009.TargetDS.payments`

A screenshot of a computer

Description automatically generated

1. List of Payment Types, respective order counts and their total payment value

SELECT distinct(payment\_type), count(order\_id) AS order\_cnt,

round(sum(payment\_value),2) AS total\_payment\_value

FROM `target-sql-361009.TargetDS.payments`

GROUP BY payment\_type

ORDER BY order\_cnt DESC

A screenshot of a computer

Description automatically generatedClearly credit card seems to be the most preferred payment type, followed by UPI.

1. Total Payment value by all orders

SELECT round(sum(payment\_value), 2) AS total\_payment\_value

FROM `target-sql-361009.TargetDS.payments`

A screenshot of a computer

Description automatically generated

1. Top 10 orders with high payment value

SELECT order\_id, payment\_value

FROM `target-sql-361009.TargetDS.payments`

ORDER BY payment\_value DESC LIMIT 10

A screenshot of a computer

Description automatically generated

1. Maximum possible payment instalment

SELECT max(payment\_installments) as max\_pmt\_installment FROM `target-sql-361009.TargetDS.payments`

A screenshot of a phone

Description automatically generated

This indicates maximum possible instalment duration is Two years.

1. Top 10 orders by payment instalment and their respective count (Number of instalments pattern)

SELECT payment\_installments AS inst\_duration, count(payment\_installments) AS inst\_cnt FROM `target-sql-361009.TargetDS.payments`

GROUP BY payment\_installments

ORDER BY inst\_cnt DESC

A screenshot of a computer

Description automatically generated

1. Top 10 orders by payment instalment and their respective count (Number of instalments pattern)

SELECT payment\_installments AS inst\_duration, count(payment\_installments) AS inst\_cnt FROM `target-sql-361009.TargetDS.payments`

GROUP BY payment\_installments

ORDER BY inst\_duration DESC

A screenshot of a computer

Description automatically generated

Based on above two data points, interestingly Customers seem to prefer/paid lower instalment than dragging it too long.

## **Orders Table**

1. Number of orders from Orders table

SELECT count(order\_id) AS order\_cnt, count(distinct(order\_id)) AS ord\_distinct\_cnt FROM `target-sql-361009.TargetDS.orders`

A screenshot of a computer

Description automatically generated

1. Number of orders and customers

SELECT count(order\_id) AS order\_cnt, count(distinct(customer\_id)) AS cust\_cnt

FROM `target-sql-361009.TargetDS.orders`

A screenshot of a computer

Description automatically generated

1. List of Order status and corresponding number of Orders

SELECT distinct(order\_status), count(order\_id) AS order\_cnt

FROM `target-sql-361009.TargetDS.orders`

GROUP BY order\_status

ORDER BY order\_cnt DESC

A screenshot of a computer

Description automatically generated

This indicates 97% of the orders are delivered & completed.

1. Delay b/w order\_delivery dates by customer & carrier

SELECT date\_diff(order\_delivered\_customer\_date, order\_delivered\_carrier\_date, DAY) as delivery\_delay,

FROM `target-sql-361009.TargetDS.orders`

ORDER BY delivery\_delay DESC LIMIT 10

A screenshot of a computer

Description automatically generated

1. Delay b/w order\_delivery dates by customer & carrier

SELECT date\_diff(CAST(order\_delivered\_customer\_date AS DATE), CAST(order\_delivered\_carrier\_date AS DATE), WEEK) as delivery\_delay\_inWeeks,

FROM `target-sql-361009.TargetDS.orders`

ORDER BY delivery\_delay\_inWeeks DESC LIMIT 10

A screenshot of a results report

Description automatically generated

1. Time duration from Order to Delivery dates

SELECT date\_diff(CAST(order\_delivered\_customer\_date AS DATE), CAST(order\_purchase\_timestamp AS DATE), WEEK) as order\_to\_delivery,

FROM `target-sql-361009.TargetDS.orders`

ORDER BY order\_to\_delivery DESC LIMIT 10

A screenshot of a results report

Description automatically generatedBased on the above queries, it’s heartening to see the ‘Order to delivery’ is 30 weeks and there around % of orders with more than 20 weeks delivery duration, which is not good.

1. Month-wise Order count

SELECT FORMAT\_DATETIME("%B", DATETIME(order\_purchase\_timestamp)) as order\_month, count(order\_id) order\_cnt

FROM `target-sql-361009.TargetDS.orders`

GROUP BY order\_month

ORDER BY order\_cnt DESC

A screenshot of a computer

Description automatically generatedAugust is the top order month. August, May and July have more number of Orders. September is the lowest number of Orders.

1. Day-wise Order count

SELECT FORMAT\_DATETIME("%a", DATETIME(order\_purchase\_timestamp)) as order\_day, count(order\_id) order\_cnt

FROM `target-sql-361009.TargetDS.orders`

GROUP BY order\_day

ORDER BY order\_cnt DESCA screenshot of a computer

Description automatically generatedSurprisingly the number of orders relatively seem to be less on week-ends compared to week-days, and Monday being the top order day.

1. Day of month with high number of Orders

SELECT FORMAT\_DATETIME("%d", DATETIME(order\_purchase\_timestamp)) as day\_of\_month, count(order\_id) order\_cnt

FROM `target-sql-361009.TargetDS.orders`

GROUP BY day\_of\_month

ORDER BY order\_cnt DESC

A screenshot of a computer

Description automatically generatedThere doesn’t seem to be any pattern in terms of which date of a month more orders are placed, like salary week or mid week or anything like that.

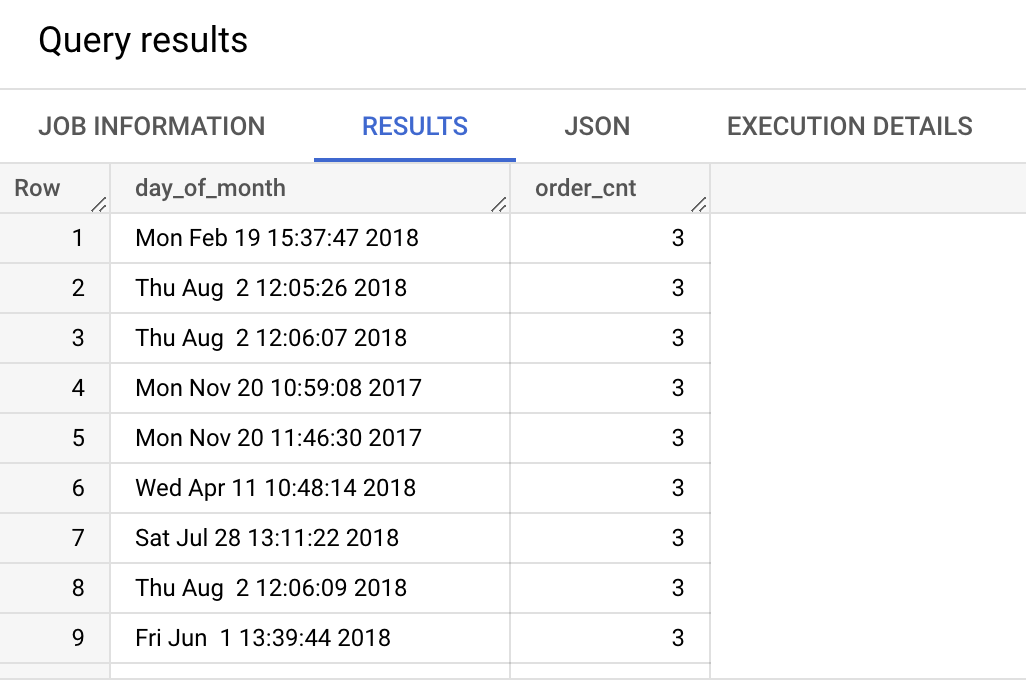
1. The days which recorded highest number of orders

SELECT FORMAT\_DATETIME("%c", DATETIME(order\_purchase\_timestamp)) as day\_of\_month, count(order\_id) order\_cnt

FROM `target-sql-361009.TargetDS.orders`

GROUP BY day\_of\_month

ORDER BY order\_cnt DESC

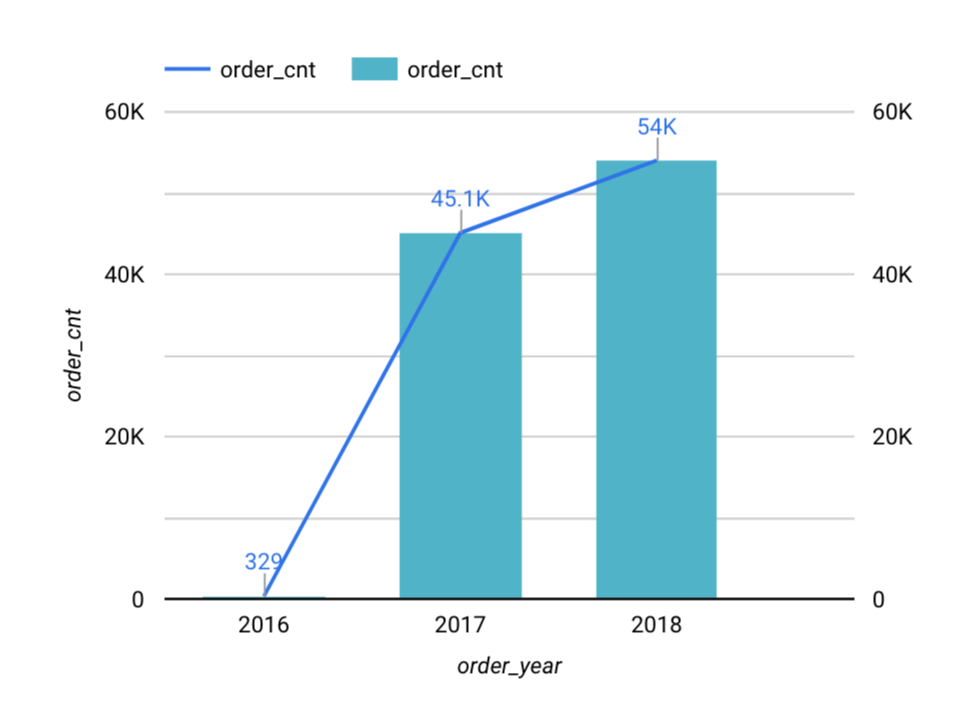
The day of the year with highest number of orders in a single day seems to be Aug 2nd.

1. Year-wise count of Orders

SELECT EXTRACT(YEAR FROM DATE (order\_purchase\_timestamp)) as order\_year, count(order\_id) order\_cnt

FROM `target-sql-361009.TargetDS.orders`

GROUP BY order\_year

ORDER BY order\_cnt DESC

This is clear indication of how sales/e-commerce has increased over the years in Brazil.

## **Reviews Table**

1. Orders with top review scores

SELECT review\_score, count(order\_id) as order\_cnt FROM `target-sql-361009.TargetDS.order\_reviews`

1. The days which recorded highest number of orders

SELECT FORMAT\_DATETIME("%c", DATETIME(order\_purchase\_timestamp)) as day\_of\_month, count(order\_id) order\_cnt

FROM `target-sql-361009.TargetDS.orders`

GROUP BY day\_of\_month

ORDER BY order\_cnt DESC

A screenshot of a data report

Description automatically generatedThe day of the year with highest number of orders in a single day seems to be Aug 2nd.

1. Year-wise count of Orders

SELECT EXTRACT(YEAR FROM DATE (order\_purchase\_timestamp)) as order\_year, count(order\_id) order\_cnt

FROM `target-sql-361009.TargetDS.orders`

GROUP BY order\_year

ORDER BY order\_cnt DESCA graph with numbers and a line

Description automatically generated

This is clear indication of how sales/e-commerce has increased over the years in Brazil.

## **Reviews Table**

1. Orders with top review scores

SELECT review\_score, count(order\_id) as order\_cnt FROM `target-sql-361009.TargetDS.order\_reviews`

GROUP BY review\_score

ORDER BY order\_cnt DESC

A screenshot of a computer

Description automatically generated57% of the total orders have got top review score and 11% has got low review score.

## **Products Table**

1. Total number of Products

SELECT count(product\_id) prod\_cnt

FROM `target-sql-361009.TargetDS.products`

A screenshot of a computer

Description automatically generated

1. Top 10 Product Categories in terms of count

SELECT product\_category, count(product\_id) prod\_cnt

FROM `target-sql-361009.TargetDS.products`

GROUP BY product\_category

ORDER BY prod\_cnt DESC LIMIT 10

A screenshot of a computer

Description automatically generatedThe top 3 product categories with more number of products are ‘bed table bath’, ‘sports’, and ‘furniture’

1. The product categories with more number of orders

SELECT product\_category, count(order\_item\_id) order\_cnt

FROM `target-sql-361009.TargetDS.products` prod

LEFT JOIN `target-sql-361009.TargetDS.order\_items` order\_item

ON prod.product\_id = order\_item.product\_id

GROUP BY product\_category

ORDER BY order\_cnt DESC LIMIT 10

A screenshot of a computer

Description automatically generated

The top 3 product categories with more number of orders are ‘bed table bath’, ‘health beauty’ and ‘sports’.

1. The product categories with less number of orders

SELECT product\_category, count(order\_item\_id) order\_cnt

FROM `target-sql-361009.TargetDS.products` prod

LEFT JOIN `target-sql-361009.TargetDS.order\_items` order\_item

ON prod.product\_id = order\_item.product\_id

GROUP BY product\_category

ORDER BY order\_cnt ASC LIMIT 10

A screenshot of a computer

Description automatically generated

Insurance and Services and Children’s Fashion Clothing are the product categories with less number of orders.

1. Top 10 Product Categories with high number of Photos

SELECT product\_category, count(product\_id) prod\_cnt, product\_photos\_qty

FROM `target-sql-361009.TargetDS.products`

GROUP BY product\_category, product\_photos\_qty

ORDER BY product\_photos\_qty DESC LIMIT 10

A screenshot of a computer

Description automatically generatedThough ‘Toys’ has only one product, the number of product images are high. Similarly we can see babies and pet products have more photos.

1. Product categories that don’t have photo images

SELECT product\_category, count(product\_id) prod\_cnt, product\_photos\_qty

FROM `target-sql-361009.TargetDS.products`

WHERE product\_photos\_qty=0

GROUP BY product\_category, product\_photos\_qty

LIMIT 10

This query has returned with zero which indicates all the products have atleast one product image.

1. Top 10 Product Categories with atleast one photo image

SELECT product\_category, count(product\_id) prod\_cnt, product\_photos\_qty

FROM `target-sql-361009.TargetDS.products`

WHERE product\_photos\_qty=1

GROUP BY product\_category, product\_photos\_qty

ORDER BY prod\_cnt DESC

LIMIT 10

A screenshot of a computer

Description automatically generated

1. Correlation b/w product order count and list of images they have

SELECT product\_category, count(order\_item\_id) order\_cnt, product\_photos\_qty

FROM `target-sql-361009.TargetDS.products` prod

JOIN `target-sql-361009.TargetDS.order\_items` order\_item

ON prod.product\_id = order\_item.product\_id

GROUP BY product\_category, product\_photos\_qty

ORDER BY order\_cnt DESC LIMIT 10

A screenshot of a data report

Description automatically generatedHaving a photo image of the product doesn’t influence on number of orders

1. Orders trend by month

SELECT FORMAT\_DATETIME("%B", DATETIME(order\_purchase\_timestamp)) as order\_month, count(order\_id) order\_cnt, customer\_state

FROM `target-sql-361009.TargetDS.orders` ord

INNER JOIN `target-sql-361009.TargetDS.customers` cust

ON ord.customer\_id = cust.customer\_id

GROUP BY order\_month, customer\_state

ORDER BY order\_cnt DESC LIMIT 12

A graph with blue and green lines

Description automatically generated

# **Insights**

The following insights can be derived from the given dataset.

Cities, States, Customers & Sellers:

* Top states with more Customers - San Paulo, Rio de janeiro
* Top states with more Sellers - San Paulo, Curitiba and Rio continue
* All the registered sellers at Target are active and they have sales
* Cities and States with more number of Sellers have recorded high number of Orders
* There is no correlation b/w the product price and high freight value

Orders:

* August is the top order month.
* August, May and July have more number of Orders.
* September has the lowest number of Orders
* Surprisingly the number of orders relatively seem to be less on week-ends compared to week-days, and Monday being the top order day.
* There doesn’t seem to be any pattern in terms of which date of a month more orders are placed, like salary week or mid week or anything like that.

Freight Value:

* High freight value have less number of orders
* Lower freight value have higher number of orders
* There are 9 orders with zero freight charges, and their total order count is 383
* 30% of orders have freight value $20 and above
* 80% of orders are from products who have freight value between 10 and 20
* 13% of orders are from products who have freight value less than 10

Payment Options:

* Preferred payment option is Credit Card with 78%
* UPI mode of payment is 17%
* Voucher holds 2% of overall payment value

Instalments:

* Maximum instalment duration is Two years
* Customers seem to prefer/paid lower instalment than dragging it too long. 80% of the instalments are within first 4 months

Delivery:

* Delivery status indicates 97% of the orders are delivered & completed
* Only 0.6% of total orders either cancelled or unavailable which is a good sign
* It’s heartening to see the ‘Order to delivery’ is 30 weeks and there around good number of orders with more than 20 weeks delivery duration, which is not good.

Reviews:

* 57% of the total orders have got top review score
* 11% of the orders have got low review score

Products:

* The top 3 product categories with more number of orders are ‘bed table bath’, ‘health beauty’ and ‘sports’.
* ‘Insurance and Services’ and ‘Children’s Fashion Clothing’ are the product categories with less number of orders.
* Though ‘Toys’ has only one product, the number of product images are high. Similarly we can see babies and pet products have more photos.
* All the products have atleast one product image.

# **Recommendations**

Based on the given dataset and above analysis, some of the recommendations are:

* States like Acre (AC), Amazonas (AM), Bahia (BA) are with less number of Customers. Incidentally Sellers count also low in those states, and there by Sales are low in these states. Target can increase the number of Sellers to attract more customers and more orders
* In such above States & Cities, Target can come up with seasonal promotions and increase Customer count
* The location/region where UPI payment is not used, probably because UPI payment type is not available OR people are not still comfortable to use that option in that location/region. Target can check and provide the option appropriately.
* The location where voucher payment type is not/less used, probably Target can give voucher promotions in that region.
* If a particular state/city is not/less using the instalment type means, probably the given duration of 2 years may not be suffice. Target can try to increase the instalment duration in those regions and see if it attracts more customers.
* Customers who make frequent purchases – recognize them with rewards, privileges
* Come up with promotions for the months like September where Sales is low
* Freight charges are really high for few orders which certainly need to be looked at reduced
* Order to delivery duration is seemingly high for few orders which certainly have to be sorted out
* ‘Insurance and Services’, ‘Children’s Fashion Clothing’, ‘PC Gamer’ are the product categories with less sales, which is an area of opportunity to improve
* There are 11% or Orders with low review scores. Target must review these scores and improve based on the given user comments.