Optimizing Target's E-commerce Operations: A Data-Driven Case Study

**Purpose and Scope:**

The purpose of this case study is to analyze Target's e-commerce operations using a comprehensive dataset. We will focus on understanding customer demographics, seller locations, payment methods, order processes, product offerings, and customer feedback. By examining these key aspects, we aim to identify insights that can inform data-driven decisions and optimize Target's online retail business. This case study will offer actionable recommendations to enhance the overall e-commerce experience for Target customers while improving operational efficiency.

**Background Information:**

Target, a retail giant, boasts a rich history dating back to 1902 and is a prominent player in the retail industry. Recent challenges include adapting to evolving consumer preferences and fierce e-commerce competition.

**Data Collection and cleaning :**

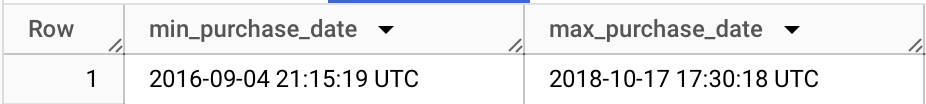
This data set has information of 100k orders from 2016 to 2018 made at Target in Brazil. Its features allows viewing an order from multiple dimensions: from order status, price, payment and freight performance to customer location, product attributes and ﬁnally reviews written by customers.

Let’s do some interesting Exploratory data analysis.

### What is the time period for which the data is provided in the "orders" table?

**SELECT MIN(order\_purchase\_timestamp) AS min\_purchase\_date, MAX(order\_purchase\_timestamp) AS max\_purchase\_date**

**FROM `target-case-study-400516.target\_store.orders`;**



From this we could see data is available between September 2016 and October 2018.

### **Are there any duplicate entries in the "sellers" table ?**

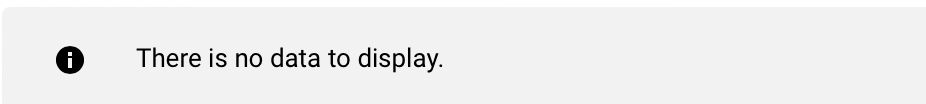
“There is no data to display” indicates there is no duplicate seller\_id and this helps in avoiding operational confusion in businesses

SELECT seller\_id, COUNT(\*)

FROM `target-case-study-400516.target\_store.sellers`

GROUP BY seller\_id

HAVING COUNT(\*) > 1;



What is the distribution of product categories ?

Understanding the product category distribution guides marketing and inventory strategies and can reveal which product categories are popular.

SELECT

product\_category,

category\_count,

ROUND((category\_count / SUM(category\_count) OVER ()) \* 100, 2) AS percentage

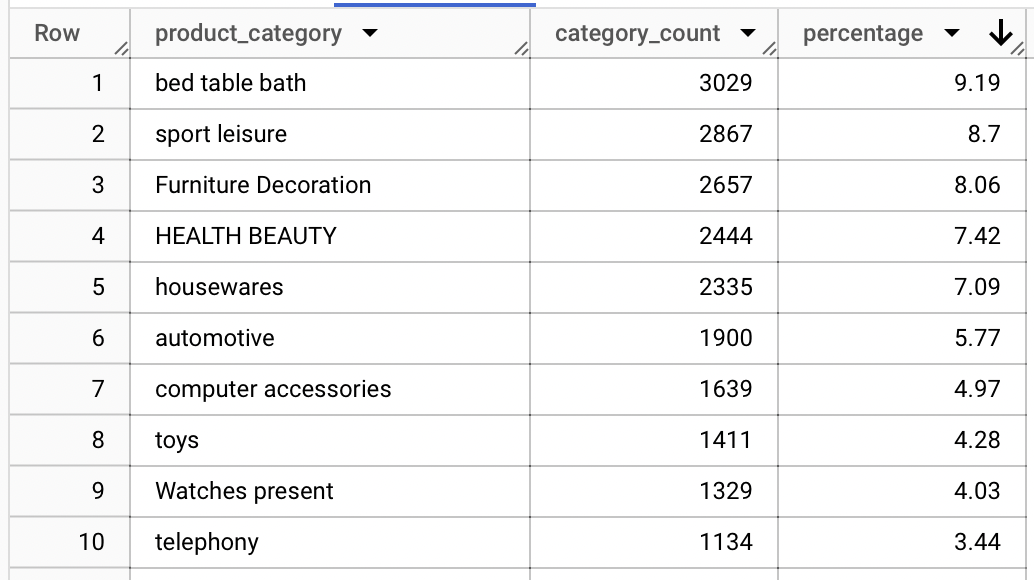
FROM

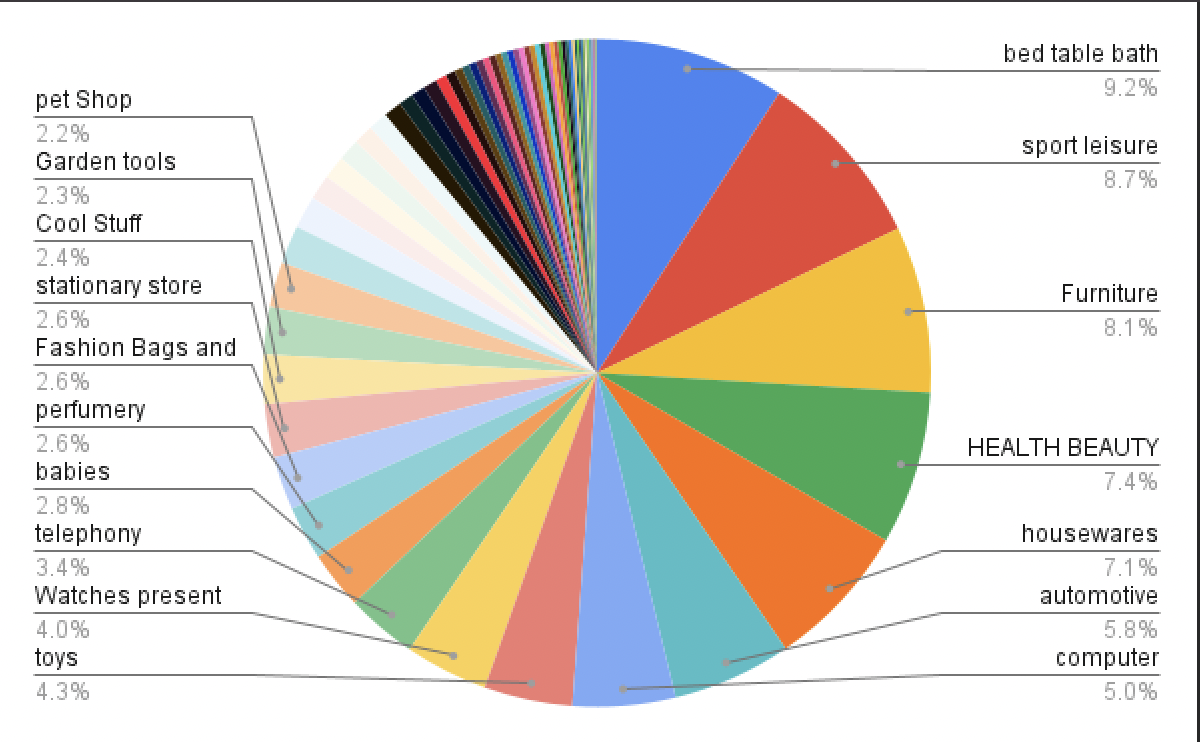
CategoryCounts

ORDER BY

category\_count DESC

LIMIT 100;





### 

### Are there any outliers in the "payment\_value" column in the "payments" table?

Outliers are data points that significantly differ from the majority of the data in a dataset

Detecting outliers is crucial for ensuring accurate analysis, as they can skew statistical measures and affect decision-making processes.

SELECT order\_id, payment\_type, payment\_value

FROM `target-case-study-400516.target\_store.payments`

WHERE payment\_value < (

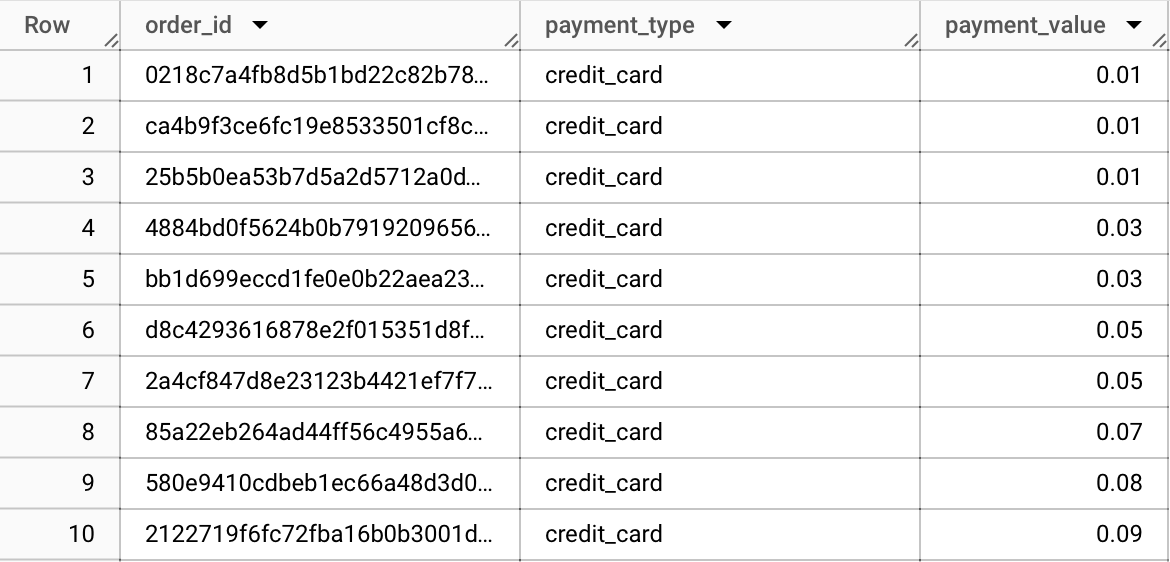
SELECT

APPROX\_QUANTILES(payment\_value, 4)[OFFSET(1)]

FROM `target-case-study-400516.target\_store.payments`

) AND payment\_type = 'credit\_card'

ORDER BY payment\_value;



SELECT order\_id,payment\_type,payment\_value

FROM `target-case-study-400516.target\_store.payments`

WHERE payment\_value > (

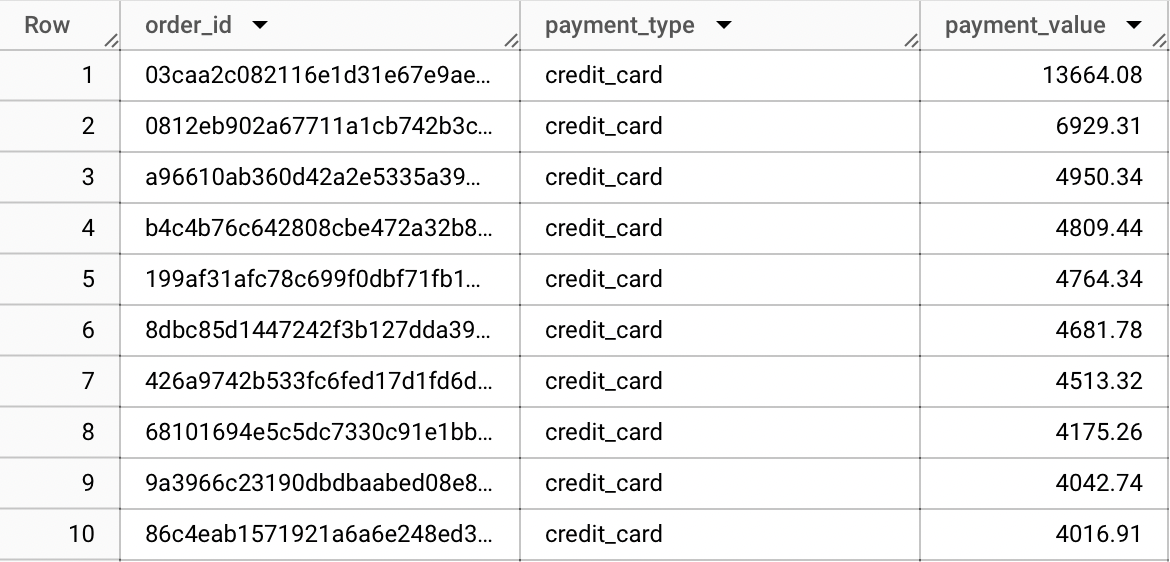
SELECT

APPROX\_QUANTILES(payment\_value, 4)[OFFSET(3)]

FROM `target-case-study-400516.target\_store.payments`

) AND payment\_type ='credit\_card'

ORDER BY payment\_value DESC;



### How many unique customers placed orders in the last quarter, and what is their geographical distribution?

SELECT customer\_city, customer\_state, COUNT(DISTINCT customer\_id) AS unique\_customers

FROM `target-case-study-400516.target\_store.customers`

WHERE customer\_id IN (

SELECT DISTINCT customer\_id

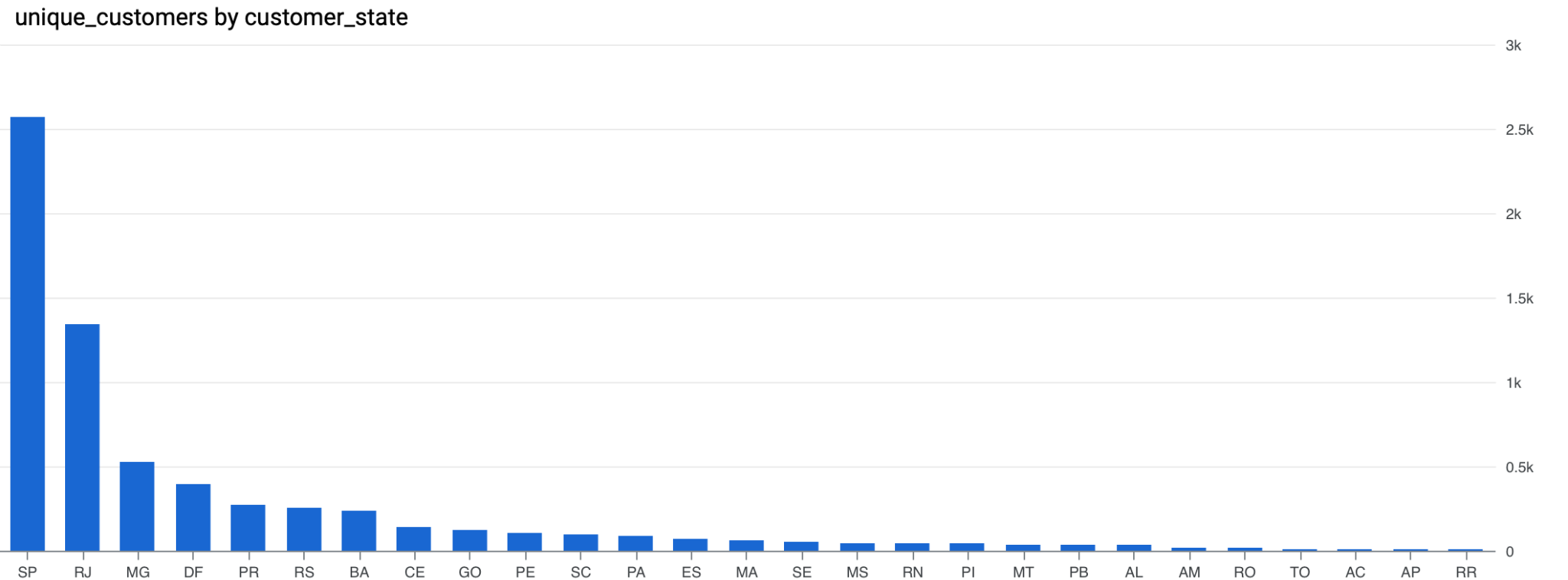
FROM `target-case-study-400516.target\_store.orders`

WHERE order\_purchase\_timestamp >= '2017-10-01' AND order\_purchase\_timestamp <= '2017-12-31'

)

GROUP BY customer\_city, customer\_state

ORDER BY uniquse\_customers desc ;



## 

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

### 

### 

### Is there a growing trend in e-commerce in Brazil? How can we describe a complete scenario?

Yes , there is a growing trend in brazil

SELECT

EXTRACT(YEAR FROM order\_purchase\_timestamp) AS year,

EXTRACT(MONTH FROM order\_purchase\_timestamp) AS month,

COUNT(1) AS num\_orders

FROM

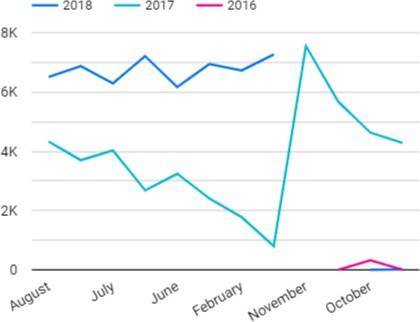
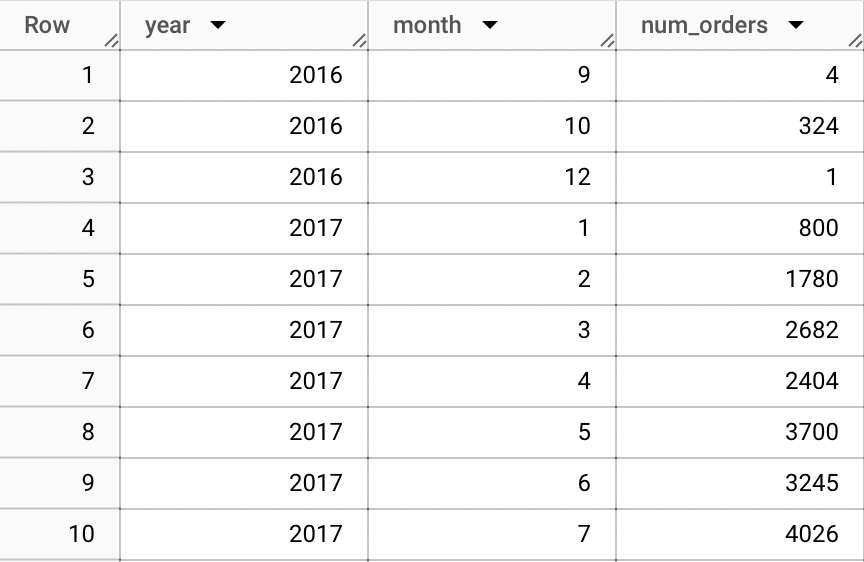
`target-case-study-400516.target\_store.orders`

GROUP BY

year, month

ORDER BY

year, month;



Can we see some seasonality with peaks at speciﬁc months?

Yes , in Brazil we can more orders toward the year end in november in december

SELECT EXTRACT(MONTH FROM order\_purchase\_timestamp) AS month, COUNT(1) AS num\_orders

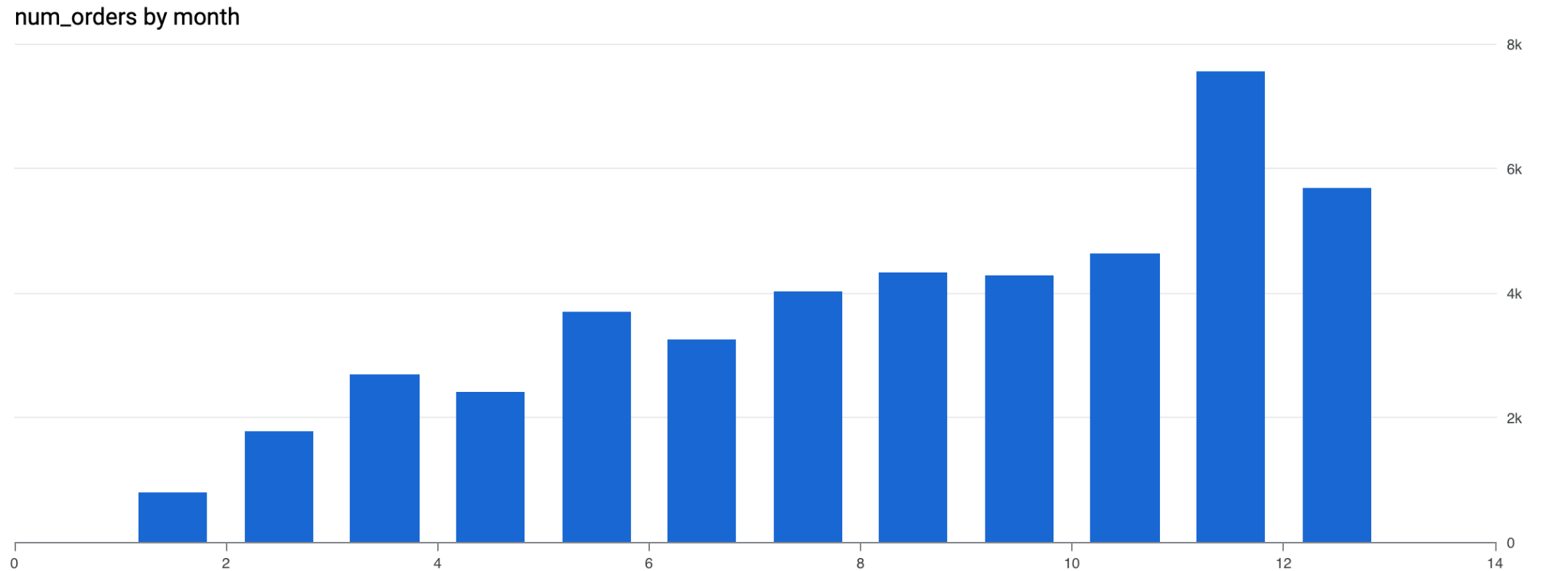
FROM `target-case-study-400516.target\_store.orders`

WHERE EXTRACT(YEAR FROM order\_purchase\_timestamp) = 2017

GROUP BY month

ORDER BY month;





What time do Brazilian customers tend to buy (Dawn, Morning, Afternoon or Night)?

Knowing when the peak ordering hours are (in the afternoon) can help in better allocation of resources. You can ensure that you have enough staff, server capacity, and logistics resources available to handle the increased demand during these hours.

WITH TimeOfDayOrders AS (

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 distinct\_orders

FROM

`target-case-study-400516.target\_store.orders`

GROUP BY

time\_of\_day

)

SELECT

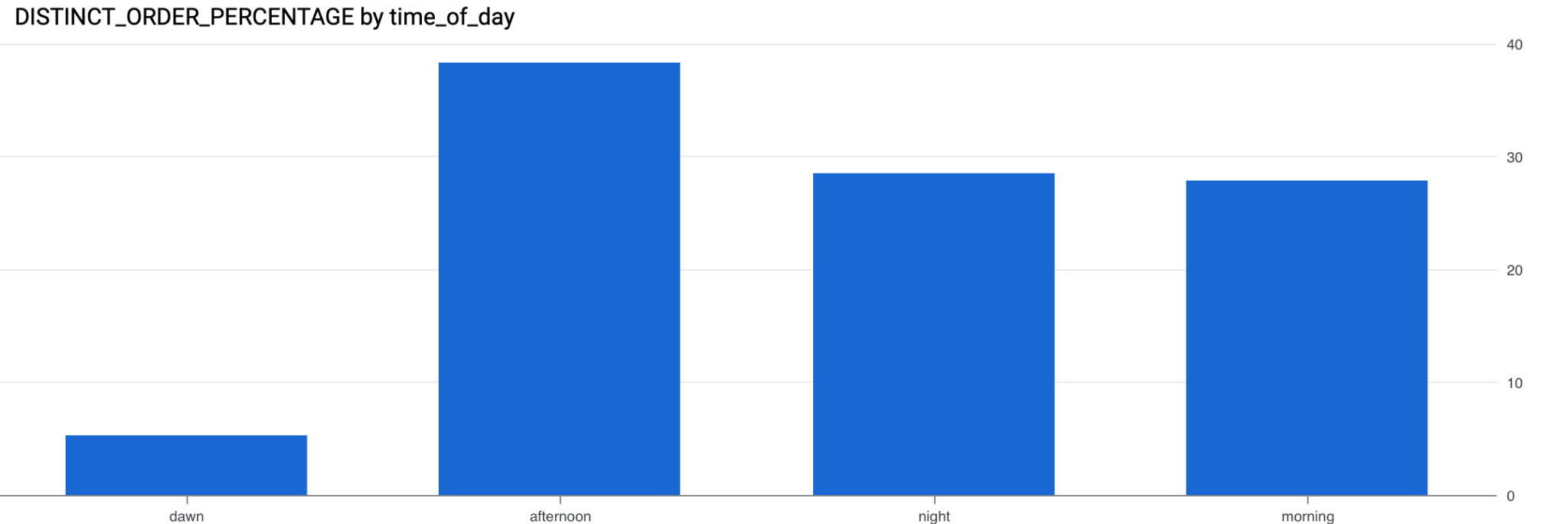
time\_of\_day,

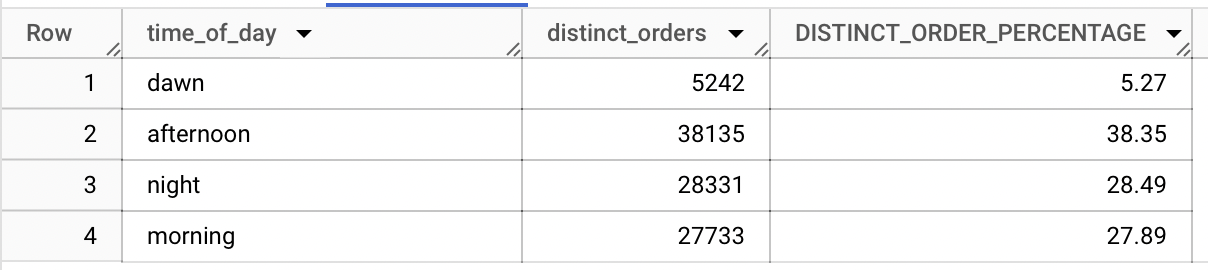
distinct\_orders,

ROUND((distinct\_orders / SUM(distinct\_orders) OVER ()) \* 100, 2) AS DISTINCT\_ORDER\_PERCENTAGE

FROM

TimeOfDayOrders;





# **Distribution of customers across the states in Brazil**

More the 50% are from top 3 states SP,RJ,MG

WITH StateCustomers AS (

SELECT

customer\_state,

COUNT(DISTINCT customer\_unique\_id) AS num\_customers

FROM

`target-case-study-400516.target\_store.customers`

GROUP BY

customer\_state

)

SELECT

customer\_state,

num\_customers,

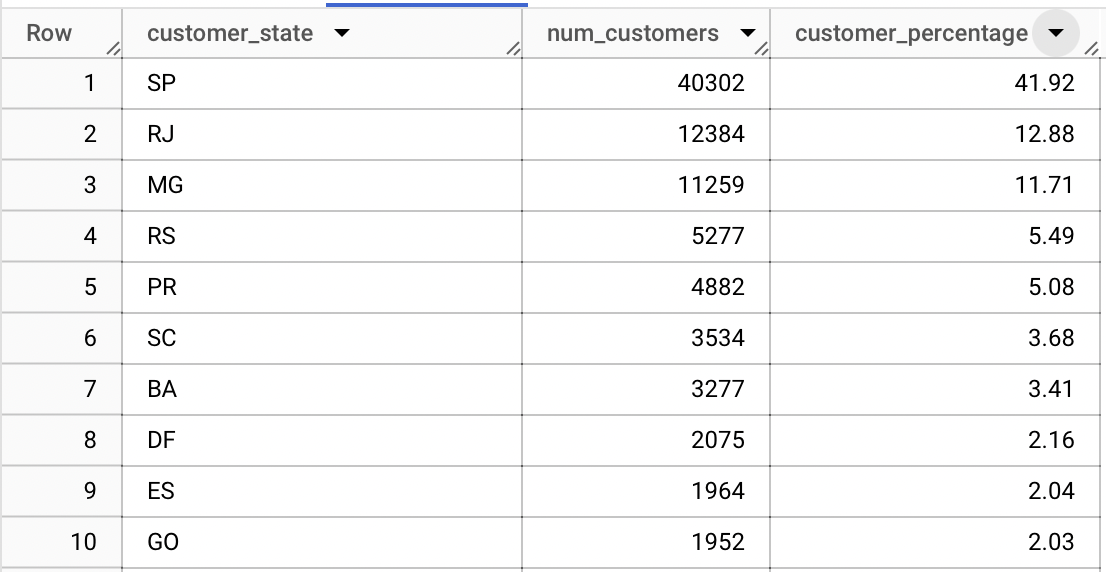
ROUND((num\_customers / SUM(num\_customers) OVER ()) \* 100, 2) AS customer\_percentage

FROM

StateCustomers

ORDER BY

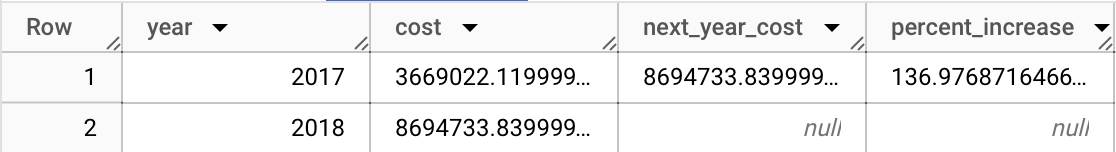
num\_customers DESC;



### Impact on Economy

# Let’s analyze the money movement by e-commerce by looking at order prices, freight and others.

We are seeing an staggering 136% growth in Brazil ecommerce from 2017 to 2018



WITH Base AS (

SELECT

EXTRACT(YEAR FROM o.order\_purchase\_timestamp) AS year,

SUM(p.payment\_value) AS cost

FROM

`target-case-study-400516.target\_store.orders` AS o

JOIN

`target-case-study-400516.target\_store.payments` AS 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

year

ORDER BY

year

),

BaseWithNextYear AS (

SELECT

\*,

LEAD(cost, 1) OVER (ORDER BY year) AS next\_year\_cost

FROM

Base

)

SELECT

\*,

(next\_year\_cost - cost) / cost \* 100 AS percent\_increase

FROM

BaseWithNextYear;

Let’s look at average order value and freight value per order

WITH cte\_table AS (

SELECT

EXTRACT(YEAR FROM o.order\_purchase\_timestamp) AS year,

EXTRACT(MONTH FROM o.order\_purchase\_timestamp) AS month,

ROUND(AVG(i.price), 2) AS rounded\_price\_per\_order,

ROUND(AVG(i.freight\_value), 2) AS rounded\_freight\_per\_order

FROM

`target-case-study-400516.target\_store.orders` AS o

INNER JOIN

`target-case-study-400516.target\_store.order\_items` AS i

ON

o.order\_id = i.order\_id

GROUP BY

year, month

ORDER BY

year, month

)

SELECT

year,

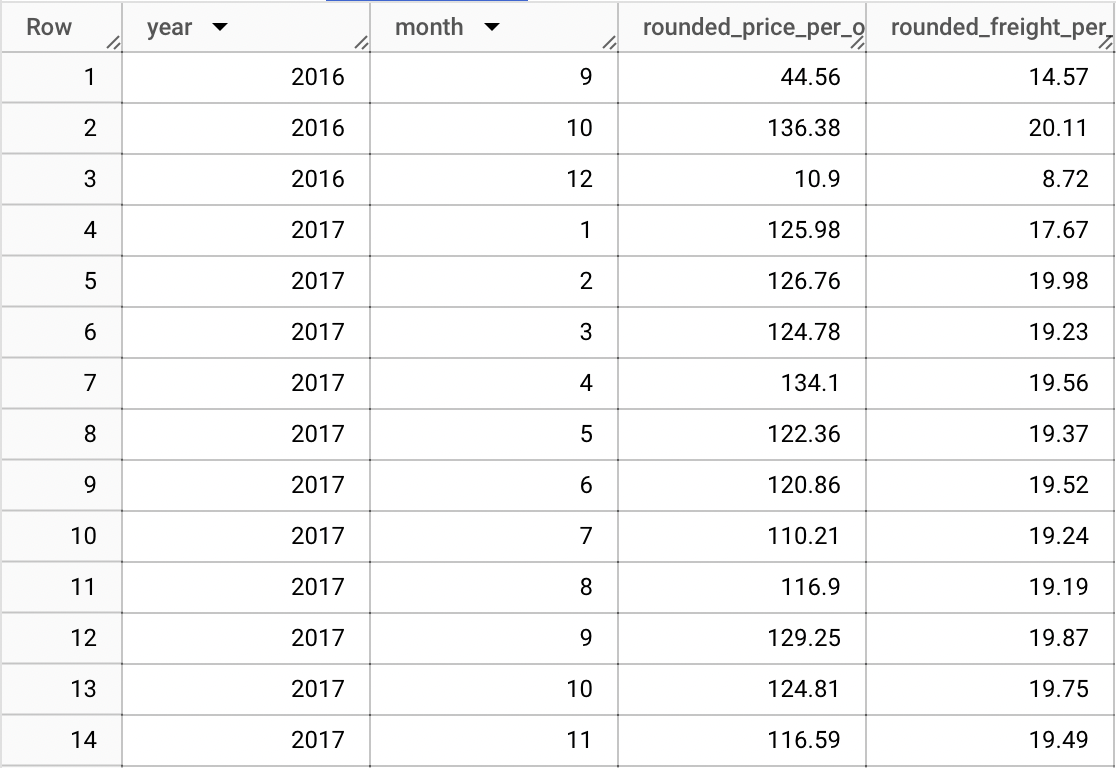
month,

rounded\_price\_per\_order,

rounded\_freight\_per\_order

FROM

cte\_table;



### Analysis on delivery time

Let’s see the time it took for delivery, and the difference between the actual delivery date and the estimated delivery date in days.

We could 90% of the orders early and we could work on predicting the estimated delivery time more accurately.

SELECT

order\_id,

TIMESTAMP\_DIFF(order\_delivered\_customer\_date, order\_purchase\_timestamp, DAY) AS time\_to\_delivery\_in\_day,

TIMESTAMP\_DIFF(order\_delivered\_customer\_date, order\_estimated\_delivery\_date, DAY) AS diff\_estimated\_delivered\_day,

CASE

WHEN TIMESTAMP\_DIFF(order\_delivered\_customer\_date, order\_estimated\_delivery\_date, DAY) < 0 THEN 'early'

WHEN TIMESTAMP\_DIFF(order\_delivered\_customer\_date, order\_estimated\_delivery\_date, DAY) > 0 THEN 'late'

ELSE 'ontime'

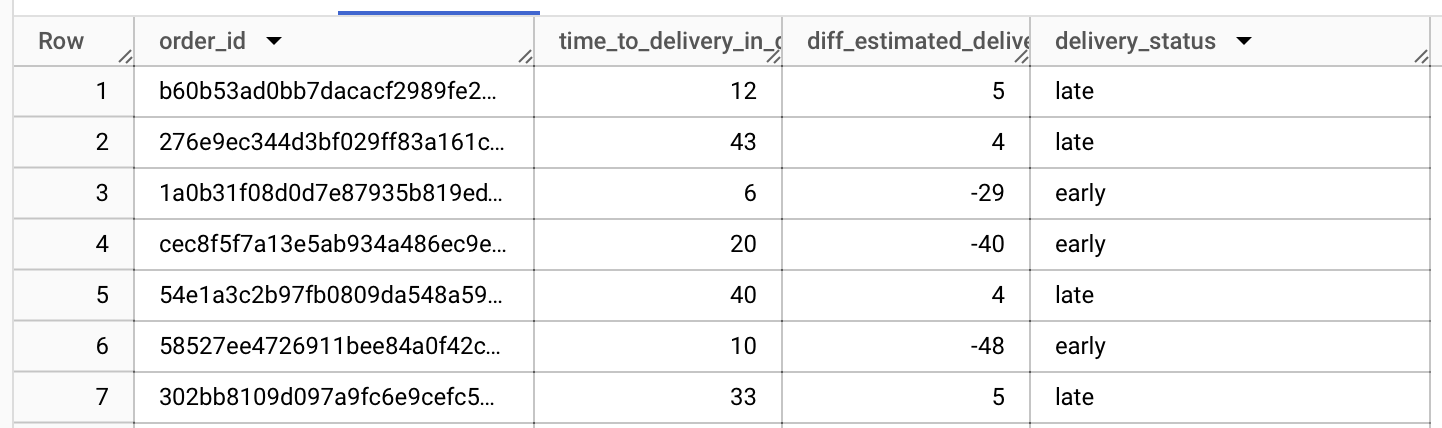
END AS delivery\_status

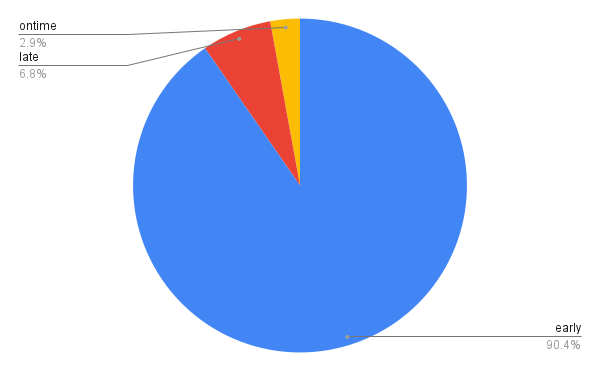
FROM

`target-case-study-400516.target\_store.orders`

WHERE

order\_status = 'delivered'





What's the popular payment type?

Brazilians prefer credits cards over other payment methods as more than ~75% payments are done via credit cards

WITH PaymentData AS (

SELECT

p.payment\_type,

EXTRACT(YEAR FROM o.order\_purchase\_timestamp) AS year,

COUNT(\*) AS num\_orders

FROM `target-case-study-400516.target\_store.orders` AS o

JOIN `target-case-study-400516.target\_store.payments` AS p

ON o.order\_id = p.order\_id

GROUP BY p.payment\_type, year

)

SELECT

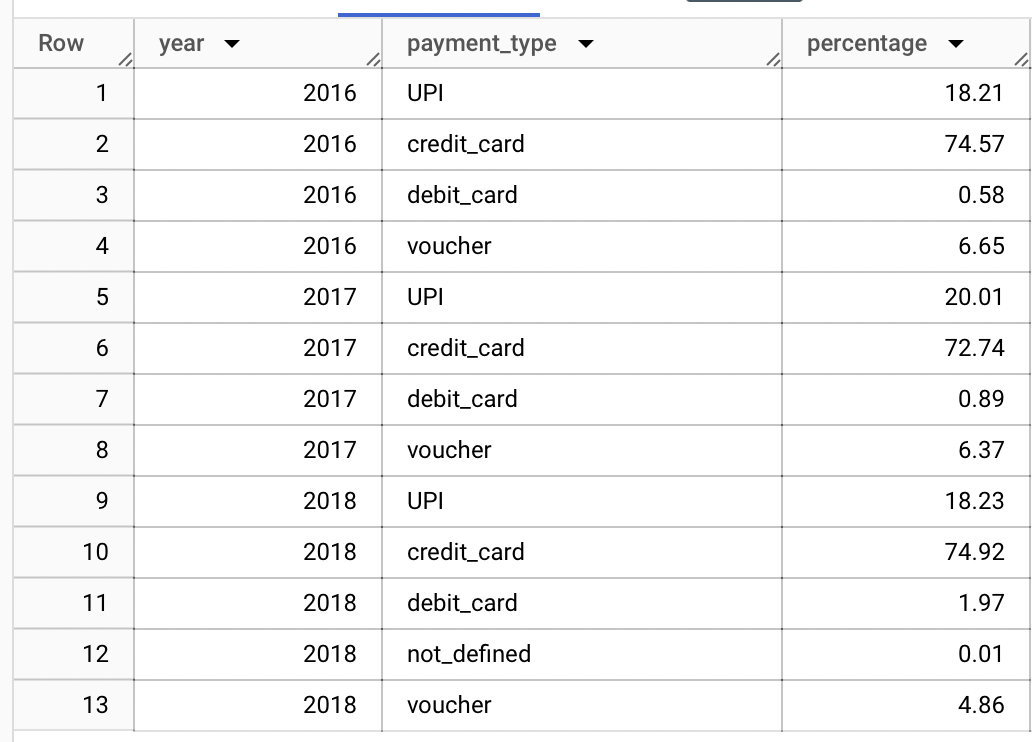
year,

payment\_type,

ROUND((num\_orders / SUM(num\_orders) OVER (PARTITION BY year)) \* 100, 2) AS percentage

FROM PaymentData

ORDER BY year, payment\_type;



### Recommendations to optimize Target's operations:

* **Improve Estimated Delivery Time Prediction:** Since the majority of orders are delivered early, it indicates an opportunity to fine-tune the estimation algorithms. This can improve customer satisfaction and help manage logistics more efficiently.
* **Enhance Marketing Strategies for Peak Months:** Capitalize on the seasonality trend, especially during November and December. Implement targeted marketing campaigns, promotions, and sufficient stock during these peak months to maximize sales.
* **Optimize Logistics and Resource Allocation:** With knowledge of the peak ordering hours in the afternoon, optimize staffing, server capacity, and logistics resources during this period. This ensures smoother operations and timely order processing.
* **Diversify Payment Options:** While credit cards are popular, consider diversifying payment options to cater to a broader audience. This may include promotions or discounts for alternative payment methods, encouraging customer adoption.
* **Improve Product Category Management:** Analyze the distribution of product categories to optimize inventory management and marketing efforts. Focus on popular categories and consider expanding or promoting lesser-known but potential categories.
* **Address Outliers in Payment Values:** Investigate and address outliers in payment values, as they can impact financial analysis. Implement strategies to handle extreme payment values more effectively..
* **Continuous Monitoring and Adaptation:** The e-commerce landscape is dynamic. Regularly monitor trends, customer preferences, and market changes. Adapt marketing, logistics, and payment strategies accordingly to stay competitive.