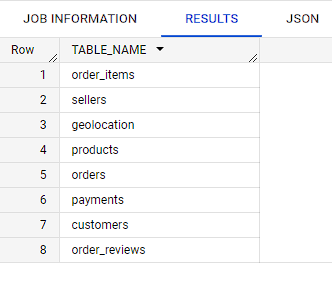
***Business Case: Target SQL Project***

1. Import the dataset and do the usual exploratory analysis steps like checking the structure and characteristics of the data set
2. Data type of columns in a table

* *Tables Present in the Database*

SELECT TABLE\_NAME

FROM Business.INFORMATION\_SCHEMA.TABLES



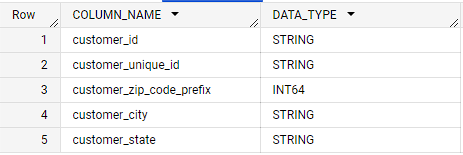
* ***Database Schema of all the tables***

**Customers Table**

SELECT COLUMN\_NAME, DATA\_TYPE

FROM Business.INFORMATION\_SCHEMA.COLUMNS

WHERE TABLE\_NAME = 'customers'

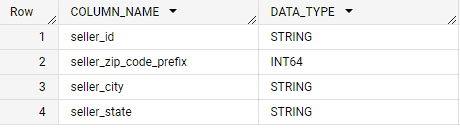


**Sellers Table**

SELECT COLUMN\_NAME, DATA\_TYPE

FROM Business.INFORMATION\_SCHEMA.COLUMNS

WHERE TABLE\_NAME = 'sellers'

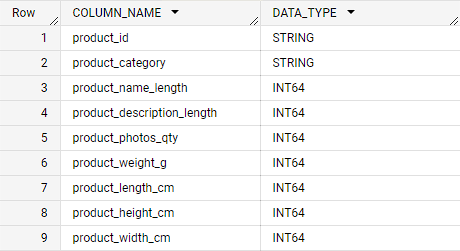
****

**Products Table**

SELECT COLUMN\_NAME, DATA\_TYPE

FROM Business.INFORMATION\_SCHEMA.COLUMNS

WHERE TABLE\_NAME = 'products'

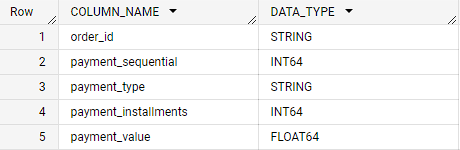
****

**Payments Table**

SELECT COLUMN\_NAME, DATA\_TYPE

FROM Business.INFORMATION\_SCHEMA.COLUMNS

WHERE TABLE\_NAME = 'payments'

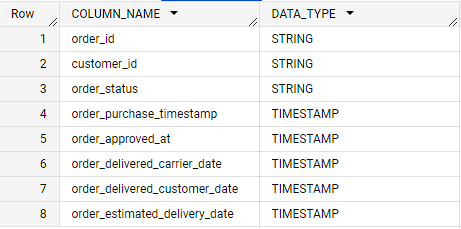


**Orders Table**

SELECT COLUMN\_NAME, DATA\_TYPE

FROM Business.INFORMATION\_SCHEMA.COLUMNS

WHERE TABLE\_NAME = 'orders'

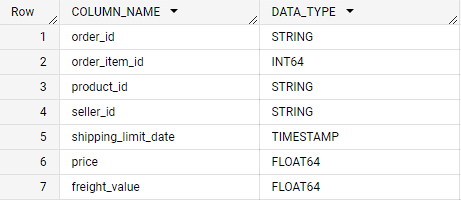


Order\_items

SELECT COLUMN\_NAME, DATA\_TYPE

FROM Business.INFORMATION\_SCHEMA.COLUMNS

WHERE TABLE\_NAME = 'order\_items'



**Insights**

* The database Schema for different tables tells about the interrelation between different tables and primary key and foreign keys

1. Time period for which the data is given

SELECT

MAX(order\_purchase\_timestamp) AS End\_date,

MIN(order\_purchase\_timestamp) AS Start\_date,

  DATE\_DIFF(MAX(order\_purchase\_timestamp), MIN(order\_purchase\_timestamp), DAY) AS Month\_Difference

FROM `Business.orders`;

**Insights**

* The difference between the End date and start date is the total time period we are concerned about.



1. Cities and states customers ordered during the given period

Distinct cities and states are present in the below table during the time period from

‘**2016-09-04 21:15:19 UTC’** to ‘ **2018-10-17 17:30:18 UTC**’ that is 772 days.



# 2. In-depth Exploration:

1. Is there a growing trend in e-commerce in Brazil? How can we describe a complete scenario? Can we see some seasonality with peaks at specific months?

SELECT

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

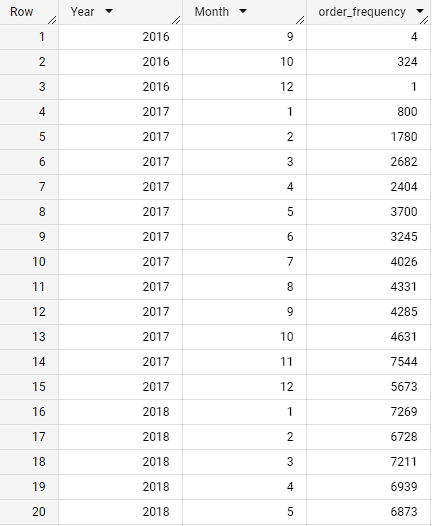
  EXTRACT(MONTH FROM (order\_purchase\_timestamp)) AS Month,

  COUNT(\*) AS order\_frequency

FROM `Business.orders`

GROUP BY 1, 2

ORDER BY Year, Month;



**Yearly order\_frequency**

SELECT

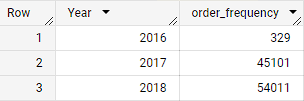
  EXTRACT(YEAR FROM order\_purchase\_timestamp) AS Year,

  COUNT(\*) as order\_frequency

FROM `Business.orders`

GROUP BY 1

ORDER BY 1



**Order Frequency monthly**

SELECT

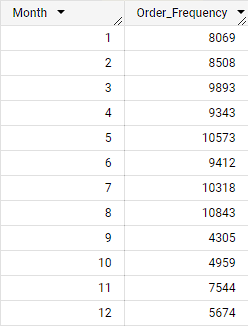
  EXTRACT(MONTH FROM (order\_purchase\_timestamp)) AS Month,

  COUNT(\*) AS Order\_Frequency

FROM `Business.orders`

GROUP BY Month

ORDER BY Month;

****

**Insights**

* **Yearly Trend**: As evident from the yearly order table there is a significant increase in orders from 2016 – 2018 which is a positive sign that people are ordering more from e-commerce sites and rapidly gaining momentum.
* **Monthly Trend:** The months from March to August have relatively higher order count compared to the rest of the months with the peak month being August where is order count is maximum this is due to the seasonal patterns as these months corresponds to autumn and Winter months in Brazil, people prefer to shop online due to reasons such as weather conditions, holiday or cultural factors and exclusive offers and deals during winter month attracts the customers to promote online shopping.

In the months from September to January e-commerce activity is low due to seasonal variations and the transition from winter to spring. Consumers might be less inclined to make online purchases during this time as they plan for year-end expenses and engage in other seasonal activities. Additionally, significant shopping events like Black Friday and Christmas lead consumers to postpone purchases and wait for discounts and promotions. Economic factors and the presence of a robust traditional retail sector with physical stores further contribute to a relatively lower proportion of online sales as consumers opt for in-person shopping experiences during the holiday season.

**Recommendation**:

* The months from March to August give an opportunity to expand Target’s presence in Brazil as these are the months where there is increased e-commerce activity, providing offers and discounts, right marketing strategies to attract customers, and having a user-friendly interface for the customers to get notified about the deals and sale prices by following these practices Target can get maximum profit and become customers favorite.
* The months from September onwards can be a crucial turning point as the e-commerce activity is low target can analyse the products that customers purchased the most and try to offer periodic discounts on them and can focus more on marketing and bring back-to-school offers so that customers continue to buy so that we can retain the customers.

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

WITH order\_hours AS (SELECT Extract(HOUR FROM order\_purchase\_timestamp) as hour,

COUNT(\*) AS order\_count

FROM `Business.orders`

GROUP BY 1)

SELECT

CASE

WHEN hour >=0 and hour<5 THEN 'Dawn'

WHEN hour>=5 and hour<12 THEN 'Morning'

WHEN hour>=12 and hour<17 THEN 'Afternoon'

WHEN hour>=17 and hour<=23 THEN 'Night'

END as time,

SUM(order\_count) as total\_count

FROM order\_hours

GROUP BY 1

ORDER BY 2



Insights:

* With the above table it is evident that the customers are mainly active in the afternoon and night period that corresponds to work hours, free time after the work hours. So, the company should focus on these hours because it is the peak buying time.

Recommendations:

* Target can focus on giving sales and discounts in order to get max sales and assist can assist this with email marketing sending notifications at these times which can benefit the company for more sales.
* In order to increase the orders in the morning and Dawn period Target can go with flash sales for 1 hour for specific items so that the order count can increase.
* Analysing customer buying patterns can positively affect the company sales.

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

* 1. Get month-on-month orders by states

SELECT C.customer\_state, FORMAT\_DATETIME("%B", DATETIME(O.order\_purchase\_timestamp)) as Month,

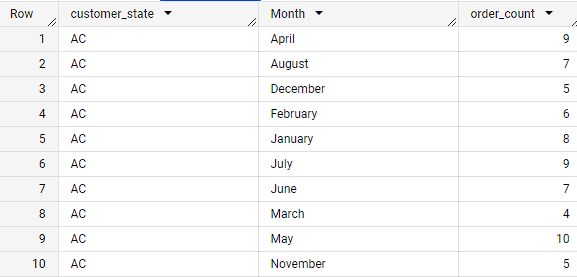
COUNT(O.order\_id) as order\_count

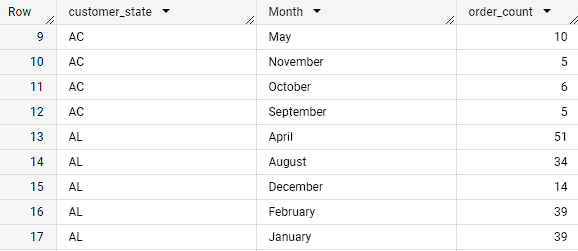
FROM `Business.customers` C

JOIN `Business.orders` O ON C.customer\_id = O.customer\_id

GROUP BY 1,2

order by 1,2





**Insights:**

* The order counts differ majorly from state to state, **SP, RJ, MG, and BA** have consistent and good numbers of order counts, and throughout the year, there is not much variation which indicates the strong presence of e-commerce in these areas.
* The months from April to August and January have higher sales due to the festival and winter season and the new year accounts for high sales in January and offers that come by with it.

**Recommendation:**

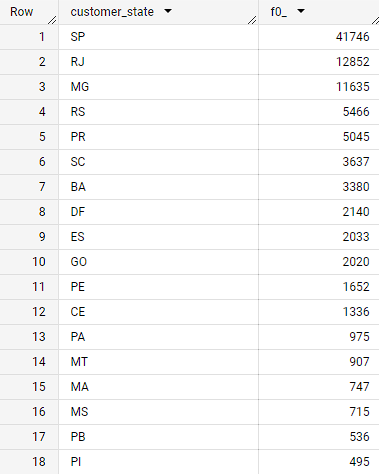
* Target should focus more on the states which have a high level of sales such as the ones mentioned in insights these are states with a strong e-commerce presence and should focus on customer satisfaction.
* Target should launch marketing campaigns in these states and analyze customer requirements by doing surveys and focusing on finding out the reason and promoting products that interest people.
* Target should focus on seasonal trends like the peak month of August which is the time when all the people buy the most therefore introducing offers, sales, and different marketing methods to gain a stronghold
  1. Distribution of customers across the states in Brazil

SELECT customer\_state , COUNT (DISTINCT customer\_id)

FROM `Business.customers`

GROUP BY 1

ORDER BY 2 DESC

****

**Insights**

* States like SP, RJ, and MG have the highest number of customers as the e-commerce presence is quite strong in these areas.
* The states of AC, AP, and RR have the lowest count and the people are not so confident in e-commerce shopping.

Recommendations:

* To increase customer count for Target in the states of RR, AP, and AL in Brazil, implement targeted marketing campaigns, establish local partnerships, offer region-specific products, enhance the customer experience, and optimize the online presence with localized content and promotions. Understanding the local market is essential.
* To increase customer count in RJ, MG, and SP, Target should prioritize improving the store experience, tailoring product selection, enhancing online presence, implementing competitive pricing and promotions, and engaging with local communities. These strategies aim to create a positive shopping environment, cater to regional preferences, optimize the online shopping experience, attract customers with attractive pricing, and build customer loyalty through community engagement.

1. **Impact on the Economy: Analyse the money movement by e-commerce by looking at order prices, freight and others.**
   1. Get a % increase in the cost of orders from 2017 to 2018 (include months between Jan to Aug only) - You can use “payment\_value” column in payments table

WITH values\_2017 AS

(SELECT round(sum(payment\_value),2) AS values\_2017

FROM `Business.payments` P

JOIN Business.orders as O

ON P.order\_id = O.order\_id

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

and EXTRACT(MONTH FROM order\_purchase\_timestamp) BETWEEN 1 and 8

ORDER BY 1 ),

values\_2018 AS

(SELECT round(sum(payment\_value),2) AS values\_2018

FROM `Business.payments` P

JOIN Business.orders as O

ON P.order\_id = O.order\_id

WHERE EXTRACT(YEAR FROM order\_purchase\_timestamp)= 2018

and EXTRACT(MONTH FROM order\_purchase\_timestamp) BETWEEN 1 and 8

ORDER BY 1)

SELECT values\_2017,values\_2018,

round(((values\_2018.values\_2018 - values\_2017.values\_2017)/values\_2017.values\_2017)\*100,2) as increase

FROM values\_2017,values\_2018;



Insights:

* There is a 136% increase in the cost of orders from previous years we can conclude that e-commerce is rapidly growing and becoming a necessity in Brazil.
* The higher cost of orders indicates increased demand and potentially higher average order values.
* E-commerce growth can be attributed to several factors, including improved internet access, increased smartphone penetration, convenience, a wider range of products, and competitive pricing. As more people embrace online shopping and businesses expand their digital presence, the e-commerce industry in Brazil is likely to continue growing.

Recommendation:

* Improve the online shopping experience with a user-friendly interface and efficient checkout process.
* Enhance customer service through prompt responses and personalized assistance.
* Expand product assortment to cater to a wider range of customer preferences.
* Strengthen marketing efforts to increase brand visibility and attract more customers.
* Invest in logistics, data analytics, and customer retention strategies to support growth and maintain competitiveness.
  1. Mean & Sum of price and freight value by a customer state

SELECT C.customer\_state,

avg(price) as average\_price,

avg(freight\_value) as average\_freight\_value,

sum(price) as sum\_price,

sum(freight\_value) as sum\_freight

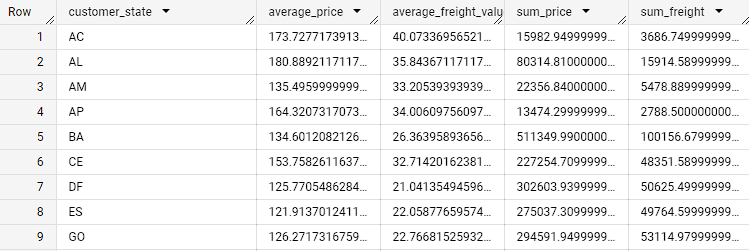
FROM `Business.customers` C

JOIN `Business.orders`O ON C.customer\_id = O.customer\_id

JOIN `Business.order\_items` OT ON O.order\_id = OT.order\_id

GROUP BY 1

ORDER BY 1;



**Insights and Recommendation:**

* Average Order Value: The average price across different customer states varies significantly, ranging from 109.65 (SP) to 191.48 (PB). Target could focus on increasing the average order value by implementing strategies such as upselling, cross-selling, and offering bundled or complementary products.
* Freight Costs: The average freight value also varies across states, with AL having the highest average at 35.84 and SP having the lowest at 15.15. Target could optimize logistics and negotiate better shipping rates to reduce freight costs, which would make their products more attractive to customers and potentially increase sales.
* Market Potential: Consider the sum of prices and sum of freight values to gauge market potential. SP has the highest sum of prices (5,202,955.05) and sum of freight (718,723.07), indicating a large market with significant revenue potential. Target could focus on expanding its presence and marketing efforts in SP to tap into this lucrative market.
* Regional Opportunities: Identify states with higher average prices and lower competition, such as PB (191.48) and AL (180.89). Target could strategically target these regions with tailored marketing campaigns and promotions to attract customers who are willing to spend more, potentially increasing sales and profitability.
* Customer Retention: Analyze customer behavior in each state to identify opportunities for improving customer retention. For instance, states like MG and RJ have significantly higher sums of prices and freight values, indicating potential loyal customer bases. Target could implement loyalty programs, personalized offers, and exceptional customer service to retain customers in these regions and encourage repeat purchases.

1. Analysis of sales, freight, and delivery time
   1. Calculate days between purchasing, delivering and estimated delivery
   2. Find time to delivery, diff estimated delivery.

SELECT order\_id,

       DATE\_DIFF(order\_delivered\_customer\_date, order\_purchase\_timestamp, DAY) AS delivery\_time,

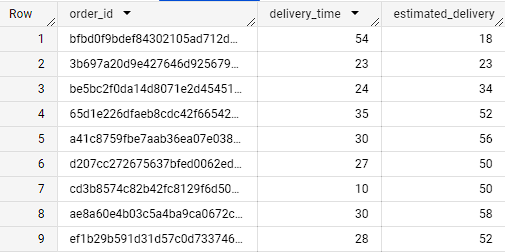
       DATE\_DIFF(order\_estimated\_delivery\_date,order\_purchase\_timestamp,DAY) as estimated\_delivery

FROM `Business.orders`

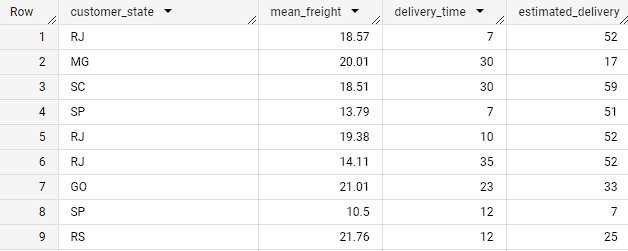
WHERE order\_delivered\_customer\_date IS NOT NULL

  AND order\_purchase\_timestamp IS NOT NULL

ORDER BY order\_purchase\_timestamp;



5.3 Group data by state, take mean of freight\_value, time\_to\_delivery, diff\_estimated\_delivery



SELECT DISTINCT c.customer\_state, ROUND(AVG(ot.freight\_value),2) as mean\_freight,

       DATE\_DIFF(order\_delivered\_customer\_date, order\_purchase\_timestamp, DAY) AS delivery\_time,

       DATE\_DIFF(order\_estimated\_delivery\_date,order\_purchase\_timestamp,DAY) as estimated\_delivery

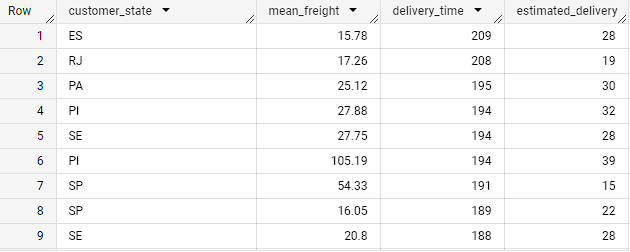
FROM `Business.customers` c

JOIN Business.orders o ON c.customer\_id = o.customer\_id

JOIN `Business.order\_items` ot ON o.order\_id= ot.order\_id

WHERE o.order\_delivered\_customer\_date IS NOT NULL

GROUP BY 1,3,4



Insights:

* The maximum time taken for delivery is 209 days whereas the estimated date is 28 which will create a bad reputation for the company this trend is followed in multiple cases the customer satisfaction will be low and could impact the image of the company.

Recommendation:

* Optimize operations: Streamline warehouse and inventory management systems, collaborate with reliable shipping partners, and establish multiple fulfillment centers strategically to improve delivery efficiency.
* Enhance last-mile delivery: Focus on improving the final leg of the delivery process by utilizing local delivery services, optimizing routes, and implementing real-time tracking and communication with customers.
* Prioritize customer experience: Offer express delivery options for faster service, maintain clear and proactive communication with customers regarding order updates and potential delays, and utilize predictive analytics to forecast demand and optimize inventory levels.

5.3 Top 5 states with highest/lowest average time to delivery

SELECT customer\_state,

        round(avg(DATE\_DIFF(order\_delivered\_customer\_date, order\_purchase\_timestamp, DAY)),2) as avg\_delivery,

FROM `Business.customers` c

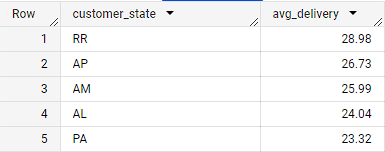
JOIN Business.orders o ON c.customer\_id = o.customer\_id

WHERE order\_delivered\_customer\_date IS NOT NULL

GROUP BY 1

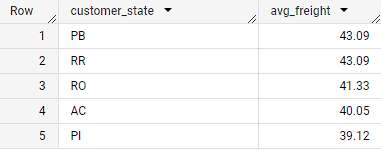
ORDER BY 2 DESC

LIMIT 5



* 1. Top 5 states with highest/lowest average freight value

Highest Freight value



SELECT customer\_state,round(avg(freight\_value),2) as avg\_freight

FROM `Business.customers` c

JOIN Business.orders o ON c.customer\_id = o.customer\_id

JOIN `Business.order\_items` ot ON o.order\_id = ot.order\_id

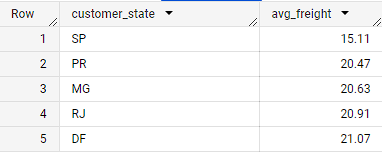
WHERE order\_delivered\_customer\_date IS NOT NULL

GROUP BY 1

ORDER BY 2 DESC

LIMIT 5

Lowest freight Value:



SELECT customer\_state,round(avg(freight\_value),2) as avg\_freight

FROM `Business.customers` c

JOIN Business.orders o ON c.customer\_id = o.customer\_id

JOIN `Business.order\_items` ot ON o.order\_id = ot.order\_id

WHERE order\_delivered\_customer\_date IS NOT NULL

GROUP BY 1

ORDER BY 2

LIMIT 5

5.5Top 5 states where delivery is really fast/ not so fast compared to estimated date

SELECT customer\_state,

        round(avg(DATE\_DIFF(order\_estimated\_delivery\_date,order\_purchase\_timestamp,DAY)),2) as avg\_estimate\_delviery

FROM `Business.customers` c

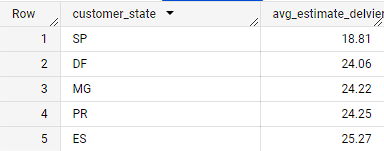
JOIN Business.orders o ON c.customer\_id = o.customer\_id

WHERE order\_estimated\_delivery\_date IS NOT NULL

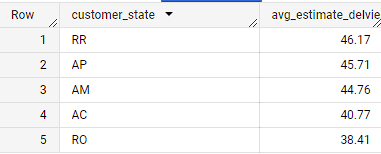
GROUP BY 1

ORDER BY 2

LIMIT 5



Not fast:



SELECT customer\_state,

        round(avg(DATE\_DIFF(order\_estimated\_delivery\_date,order\_purchase\_timestamp,DAY)),2) as avg\_estimate\_delviery

FROM `Business.customers` c

JOIN Business.orders o ON c.customer\_id = o.customer\_id

WHERE order\_estimated\_delivery\_date IS NOT NULL

GROUP BY 1

ORDER BY 2 DESC

LIMIT 5

1. **Payment type analysis**

SELECT EXTRACT(YEAR FROM O.order\_purchase\_timestamp) as YEAR,

       EXTRACT(Month FROM O.order\_purchase\_timestamp) as Month,

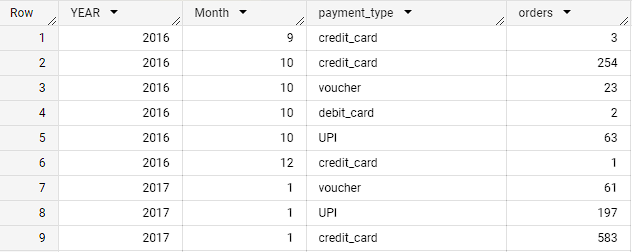
       payment\_type,count(O.order\_id) as orders

FROM `Business.orders` O

JOIN Business.payments P ON O.order\_id = P.order\_id

GROUP BY 1,2,3

ORDER BY 1,2

****

**Insights:**

* We can observe that year by year credit card payments are increasing followed by UPI payments as it is easy and can be used for faster transactions.
* Debit card payments are not much popular as there is a very low probability that customer uses a debit card
* Vouchers are increasingly used by people more and more

**Recommendation:**

* To retain credit card customers Target can provide offers on credit cards frequently so that the credit card payments can increase more and tie up with banks to issue Target Credit cards which can offer cash backs
* Simplify the payment process and remove any unnecessary steps or friction points to make it easy and convenient for customers to use debit card and UPI payments.
* Promote the benefits of debit card and UPI payments, highlighting their convenience, security, and instant transfer capabilities, and offer incentives such as cashbacks, discounts, or rewards to encourage customers to choose these payment options.
* Optimize the mobile experience, collaborate with banks and UPI providers, leverage social media and email marketing, and provide excellent customer support to further enhance the adoption of debit card and UPI payments for Target e-commerce.

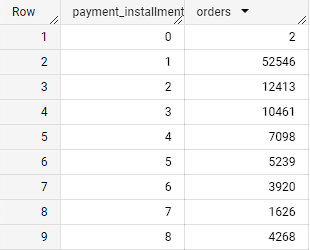
**6.2** Count of orders based on the no. of payment installments

SELECT payment\_installments,COUNT(order\_id) as orders

FROM `Business.payments`

GROUP BY 1

ORDER BY 1,2



**Insights :**

* The majority share is taken by one-time payments followed by 2 installments and 3 installments payments.
* Instalments 4 to 8 are considerably less as people don’t want long-term Emi.

**Recommendation:**

* Offer flexible payment options such as one-time payments and installment plans at the checkout to cater to different customer preferences and financial situations.
* Clearly communicate the pricing, financing terms, and any associated fees or interest rates for installment plans to ensure transparency and build trust with customers.
* Provide exclusive promotions, discounts, or incentives for customers who choose installment plans, encouraging them to opt for this payment method.
* Empower customers by allowing them to customize their payment schedules and choose payment options that suit their needs.
* By providing flexible payment options and incentives, you can increase e-commerce sales for Target and attract customers who may have been hesitant to make larger purchases.