**Project Target-SQL**

About Target:

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.Analyzing this business case helps us gain valuable insights into Target’s operations in Brazil

Dataset:

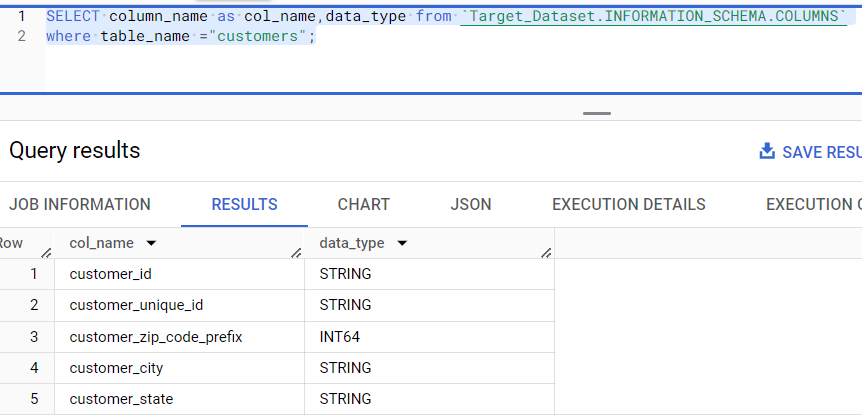
<https://drive.google.com/drive/folders/1TGEc66YKbD443nslRi1bWgVd238gJCnb>

**Exploratory Analysis**:

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

SELECT column\_name as col\_name,data\_type from `Target\_Dataset.INFORMATION\_SCHEMA.COLUMNS`

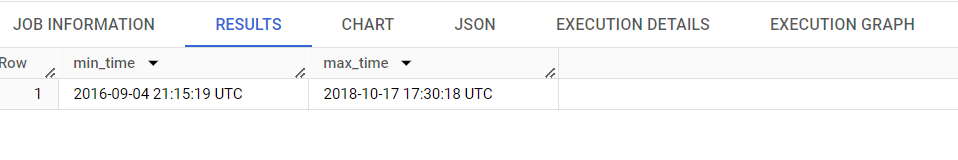
where table\_name ="customers";



2) Get the time range between which the orders were placed.

select min(order\_purchase\_timestamp) as min\_time ,max(order\_purchase\_timestamp) as max\_time

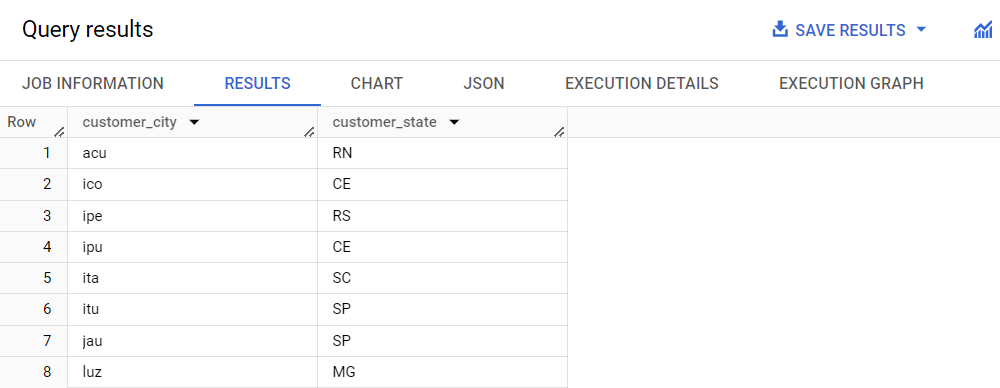
from `Target\_Dataset.orders`;



3) Count the Cities & States of customers who ordered during the given period

select c.customer\_city,c.customer\_state from `Target\_Dataset.customers` c inner join `Target\_Dataset.orders` o

on c.customer\_id=o.customer\_id group by 1,2;



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

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

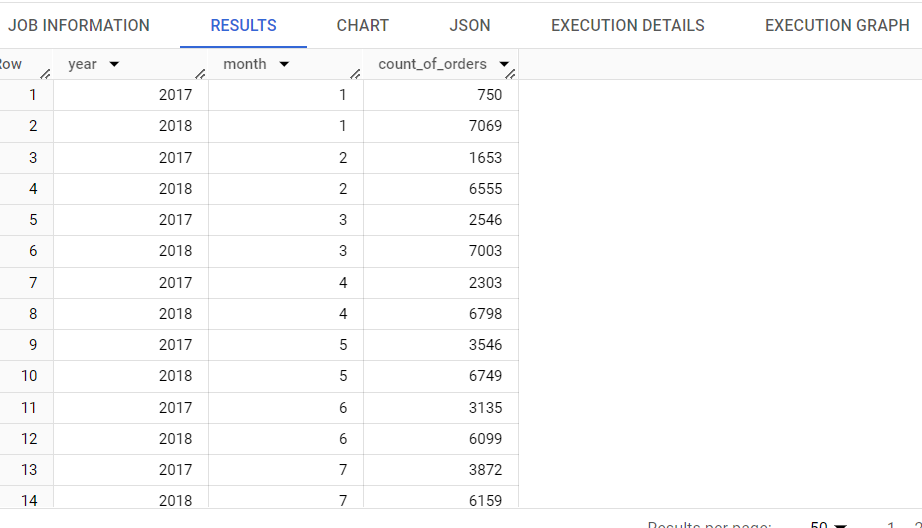
extract(month from order\_purchase\_timestamp) as month,

count(distinct(order\_id)) as count\_of\_orders

 from `Target\_Dataset.orders`

 where order\_status="delivered"

 group by 1,2 order by 2,1;



Alternate way:

You can also find out the maximum orders for every year and every month where the sales were at peak.

create view Target\_Dataset.seasonality as (select extract(year from order\_purchase\_timestamp ) as year,

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

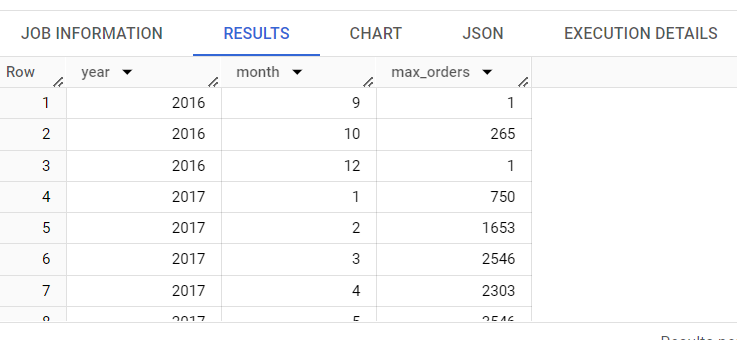
count(distinct(order\_id)) as count\_of\_orders

 from `Target\_Dataset.orders`

 where order\_status="delivered"

 group by 1,2 order by 2,1);

select year,month,max(count\_of\_orders) as max\_orders from Target\_Dataset.seasonality group by year,month order by 1,2;



5) 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

select sum(case when hour\_ between 0 and 6 then orders end) as Dawn,

sum(case when hour\_ between 7 and 12 then orders end) as Mornings,

sum(case when hour\_ between 13 and 18 then orders end)as Afternoon,

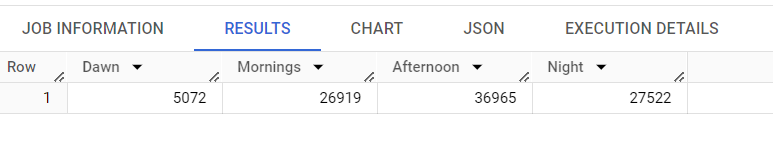
sum(case when hour\_ between 19 and 23 then orders end)as Night

from

(select extract(hour from order\_purchase\_timestamp) as hour\_,

count(distinct(order\_id)) as orders from `Target\_Dataset.orders`

where order\_status="delivered" group by 1)



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

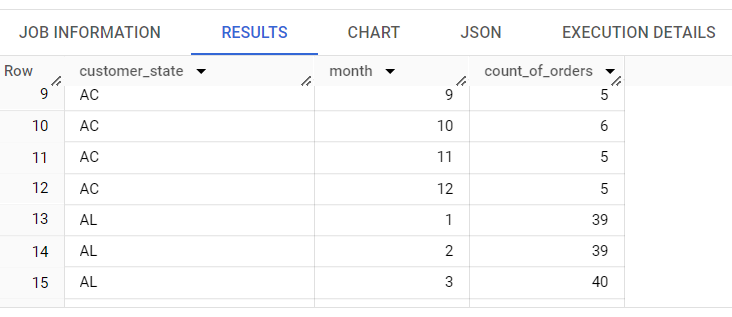
1)Month on month no. of orders placed in each state.

select  c.customer\_state,extract(month from o.order\_purchase\_timestamp) as month,

count(distinct(o.order\_id)) as count\_of\_orders

 from`Target\_Dataset.customers` c inner join `Target\_Dataset.orders` o

on c.customer\_id=o.customer\_id group by 1,2 order by 1,2;

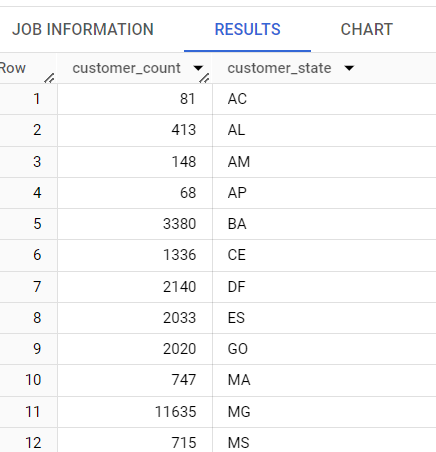


.

2) How are the customers distributed across all the states?

select count(customer\_unique\_id) as customer\_count,customer\_state from `Target\_Dataset.customers`

group by 2 order by 2,1



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)

with pay as

(select extract(year from o.order\_purchase\_timestamp)as year,round(sum(p.payment\_value),0) as payment from `Target\_Dataset.payments`p inner join `Target\_Dataset.orders`o

on p.order\_id=o.order\_id

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

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

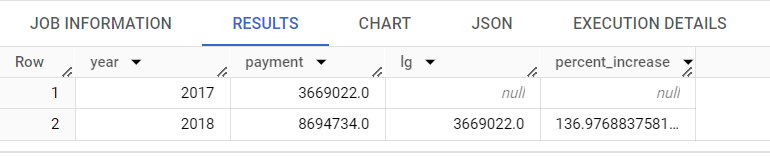
group by 1

),

lagger as(

select \*,lag(payment,1) over(order by year asc) as lg from pay)

select \*,((payment-lg)/lg)\*100 as percent\_increase from lagger



2) Calculate the Total & Average value of order price for each state.

select c.customer\_state,round(sum(oi.price),1) as total\_sum,round(avg(oi.price),1) as avg\_price

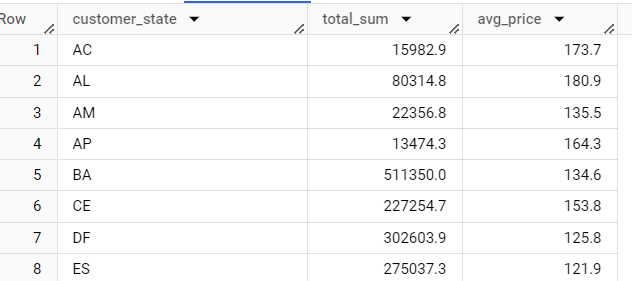
 from

 `Target\_Dataset.customers` c inner join `Target\_Dataset.orders`o

on c.customer\_id=o.customer\_id inner join

`Target\_Dataset.order\_items` oi on o.order\_id=oi.order\_id

group by c.customer\_state order by 1,2



3) Calculate the Total & Average value of order freight for each state.

select c.customer\_state,round(sum(oi.freight\_value),1) as total\_sum,round(avg(oi.freight\_value),1) as avg\_price

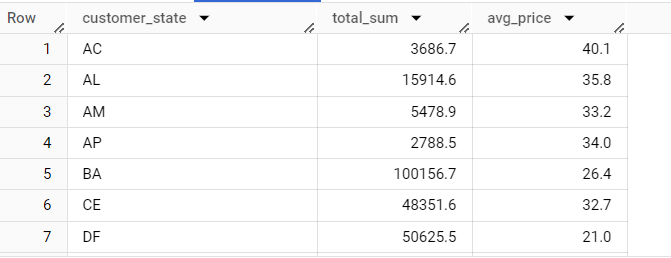
 from

 `Target\_Dataset.customers` c inner join `Target\_Dataset.orders`o

on c.customer\_id=o.customer\_id inner join

`Target\_Dataset.order\_items` oi on o.order\_id=oi.order\_id

group by c.customer\_state order by 1,2



**Analysis based on sales, freight and delivery time:**

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.

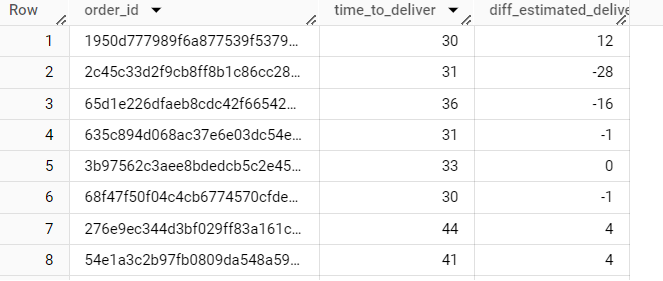
SELECT

    order\_id,

    DATE\_DIFF(order\_delivered\_customer\_date, TIMESTAMP\_TRUNC(order\_purchase\_timestamp, DAY), DAY) AS time\_to\_deliver,

    DATE\_DIFF(order\_delivered\_customer\_date,order\_estimated\_delivery\_date,day) AS diff\_estimated\_delivery

FROM Target\_Dataset.orders`;



2) Find out the top 5 states with the highest & lowest average freight value.

    select c.customer\_state,avg(freight\_value) as max\_freight

    from `Target\_Dataset.customers` c inner join `Target\_Dataset.orders` o on

    c.customer\_id=o.customer\_id inner join `Target\_Dataset.order\_items` oi on

    o.order\_id=oi.order\_id group by c.customer\_state order by avg(freight\_value) desc

    limit 5

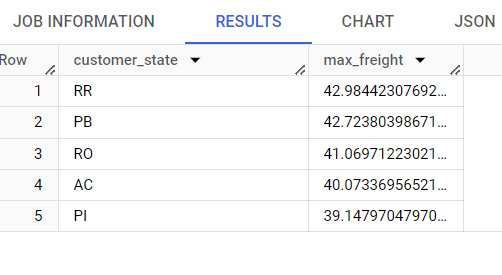
    select c.customer\_state,avg(freight\_value) as min\_freight

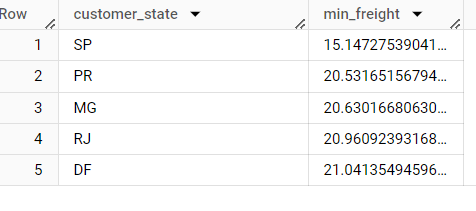
    from `Target\_Dataset.customers` c inner join `Target\_Dataset.orders` o on

    c.customer\_id=o.customer\_id inner join `Target\_Dataset.order\_items` oi on

    o.order\_id=oi.order\_id group by c.customer\_state order by avg(freight\_value) asc

    limit 5





**Analysis based on the payments:**

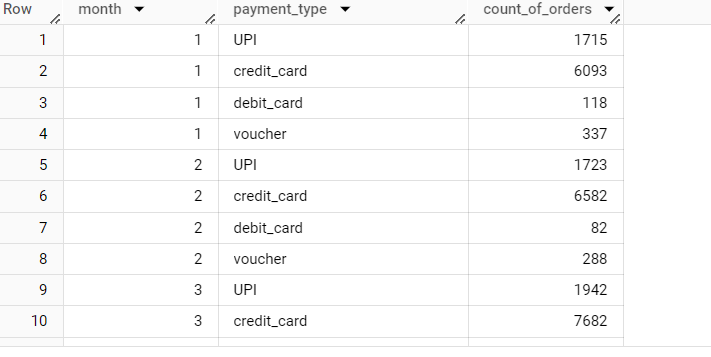
1) Find the month on month no. of orders placed using different payment types.

  select extract(month from o.order\_purchase\_timestamp) as month,p.payment\_type,

  count(distinct(o.order\_id)) as count\_of\_orders

  from`Target\_Dataset.payments` p inner join `Target\_Dataset.orders` o

  on p.order\_id=o.order\_id group by 1,2 order by 1,2;



Observations made:

1)Target had improved its sales from the year 2016 to 2017,however there was a slight decrease in its performance in the year 2018.

2)Brazilian customers are more active during afternoon as they place huge number of orders.So improving marketing strategies,offers during these hours will have a positive effect on the company’s performance.

3)It is observed that target takes a lot of time to deliver products to its customers. Improving its delivery time will increase its customer base as they would prefer target over any other retail store with more delivery time.

4) It is also observed that credit card sales/payments play a major role .So connecting with more credit card companies and providing credit card offers/options have a higher chance of increasing its sales.