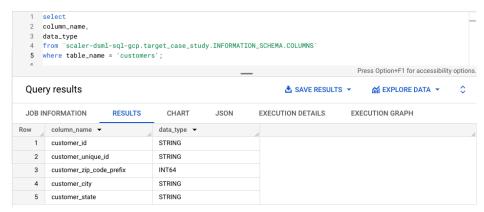
# 1) Import the dataset and do usual exploratory analysis steps like checking the structure & characteristics of the dataset:

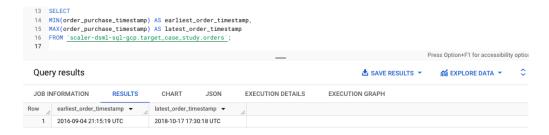
 I have imported all the 8 tables from the target data sets in my BigQuery platform and started analysing all the tables and its characteristics generally.

#### a) Data type of all columns in the "customers" table



- I just verified all the data types of each column from the customers table by fetching information from the table.
- Except customer\_zip\_code\_prefix column all the other columns are string data type.

#### b)Get the time range between which the orders were placed.



As per the requirement, I have analysed the range of orders placed during a given interval of time period.

I used max(), min() functions to fetch range.

Minimum timestamp of order were placed on 2016-09-04 21:15:19 UTC Maximum timestamp of order were placed on 2018-10-17 17:30:18 UTC

#### c)Count the Cities & States of customers who ordered during the given period.



 It results totally 4119 cities and 27 states of customers were placed order during the given period

2) In-depth Exploration:

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



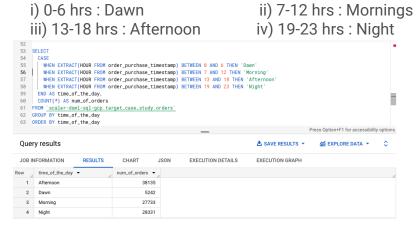
There are drastic changes in the number of orders placed between 2016 and 2017,18.

b)Can we see some kind of monthly seasonality in terms of the no. of orders being placed?



• It results number of orders placed month wise

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



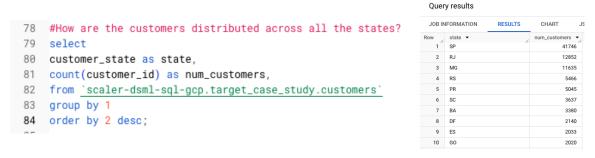
It results orders been placed different times in a day

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

a)Get the month on month no. of orders placed in each state.

```
66 #Get the month on month no. of orders placed in each state.
                                                                                 JOB INFORMATION
                                                                                                RESULTS
                                                                                                         CHART
67 select
                                                                                              month •
68 c.customer_state,
69 extract(month from order_purchase_timestamp) as month,
                                                                                                             4632
70 count(1) as num_orders
                                                                                                             4381
                                                                                                             4104
71 from `scaler-dsml-sql-gcp.target_case_study.orders` o
                                                                                                             4047
72 INNER JOIN <u>`scaler-dsml-sql-gcp.target_case_study.customers`</u> c
                                                                                                             3967
73 ON o.customer_id = c.customer_id
74 group by 1,2
75 order by 3 desc;
                                                                                10
                                                                                                             2357
```

#### b) How are the customers distributed across all the states?



4.Impact on Economy: Analyze the money movement by e-commerce by looking at order prices, freight and others.

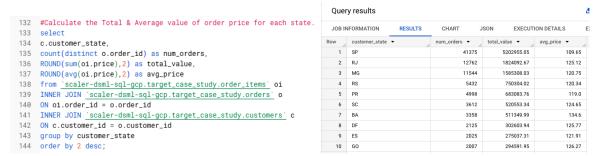
a)Get the % increase in the cost of orders from year 2017 to 2018 (include months between Jan to Aug only).

```
LEAD(cost) OVER (ORDER BY year) AS next_year_cost,
ROUND(((LEAD(cost) OVER (ORDER BY year) - cost) / cost) * 100, 2) AS percent_increase
         scaler-dsml-sql-gcp.target_case_study.orders' o
INNER JOIN 'scaler-dsml-sql-gcp.target_case_study.payments' p ON o.order_id = p.order_id
            ME.

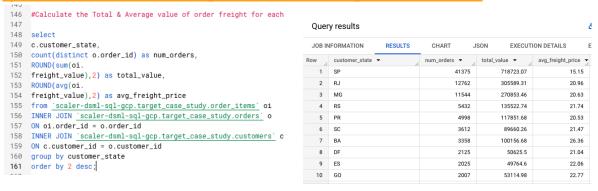
EXTRACT(YEAR FROM order_purchase_timestamp) BETWEEN 2017 AND 2018

AND EXTRACT(MONTH FROM order_purchase_timestamp) BETWEEN 1 AND 8
     GROUP BY
| EXTRACT(YEAR FROM order_purchase_timestamp)
| AS yearly_costs;
      Query results
      JOB INFORMATION
                                                  RESULTS
                                                                              CHART
                                                                                                      JSON
                                                                                                                             EXECUTION DETAILS
   Row
                   year -
                                                                                       next_year_cost ▼
                                                                                                                         percent_increase -
                                       2018
                                                     8694733.839999...
          1
                                                                                                                                              null
          2
                                       2017
                                                     3669022.119999...
                                                                                       8694733.839999...
                                                                                                                                          136.98
```

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



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



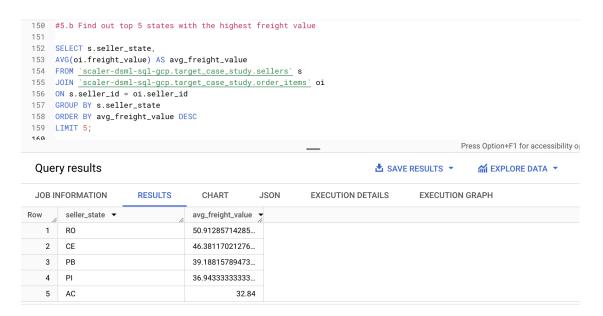
#### 5. Analysis based on sales, freight and delivery time.

<u>a)Find the no. of days taken to deliver each order from the order's purchase date as</u> delivery time.

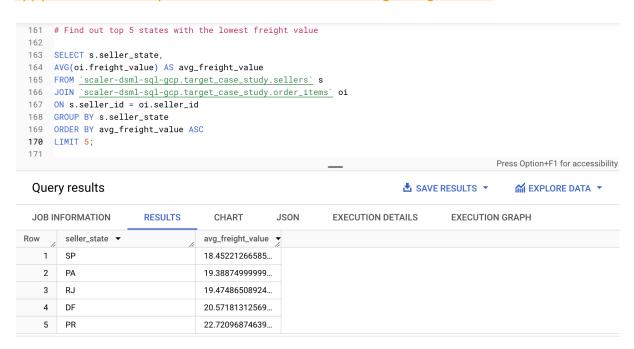
Also, calculate the difference (in days) between the estimated & actual delivery date of an order.



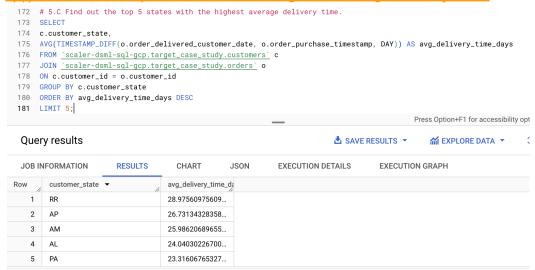
### b)(i) Find out the top 5 states with the highest average freight value.



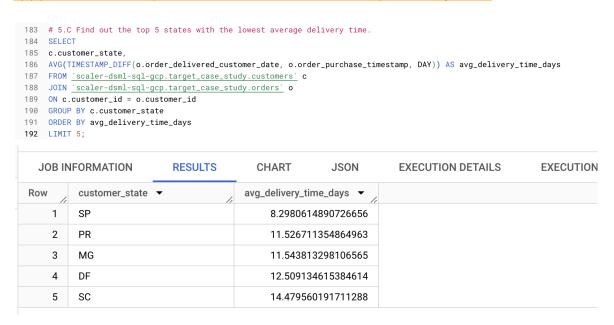
#### b)(ii) Find out the top 5 states with the lowest average freight value.



#### c)(i) Find out the top 5 states with the highest average delivery time.



#### c)(ii) Find out the top 5 states with the lowest average delivery time.



# d) Find out the top 5 states where the order delivery is really fast as compared to the estimated date of delivery.

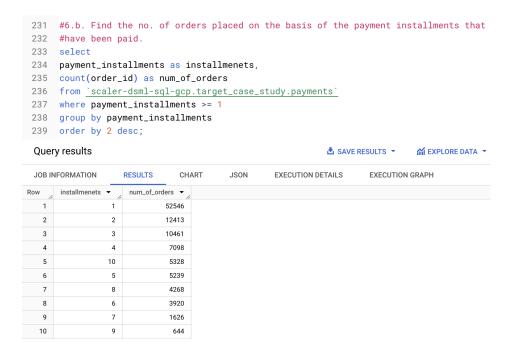
```
200 #5.d. Find out the top 5 states where the order delivery is really fast as compared
201 # to the estimated date of delivery.
202
203 SELECT
204 customer_state AS state,
205 ROUND(SUM(TIMESTAMP_DIFF(order_delivered_customer_date, order_purchase_timestamp,
206 DAY))/COUNT(ORDER_ID), 2) AS average_time_for_delivery,
207 ROUND(SUM(TIMESTAMP_DIFF(order_estimated_delivery_date, order_purchase_timestamp,
208 DAY))/COUNT(ORDER_ID), 2) AS average_est_del_time,
209 FROM \underline{\text{`scaler-dsml-sql-gcp.target\_case\_study.orders'}} o
210 INNER JOIN <u>`scaler-dsml-sql-gcp.target_case_study.customers`</u> c
211 ON o.customer_id=c.customer_id
212 WHERE order_status='delivered'
213 GROUP BY customer_state
214 ORDER BY (average_time_for_delivery-average_est_del_time)
215 limit 5;
```

B INFORMATION	N F	RESULTS	CHA	ART JSON EXE	ECUTION DETAILS	EXECUTION GRAPH
year ▼	4	month ▼	/	payment_type ▼	num_orders ▼	
1	2016		9	credit_card	3	
2	2016		10	UPI	63	
3	2016		10	credit_card	254	
4	2016		10	debit_card	2	
5	2016		10	voucher	23	
6	2016		12	credit_card	1	
7	2017		1	UPI	197	
8	2017		1	credit_card	583	
9	2017		1	debit_card	9	
10	2017		1	voucher	61	

#### 6). Analysis based on the payments:

```
a)Find the month on month no. of orders placed using different payment types.
  217 #6.a. Find the month on month no. of orders placed using different payment types.
  218
  219 SELECT
  220 EXTRACT(YEAR FROM o.order_purchase_timestamp) AS year,
  221 EXTRACT(MONTH FROM o.order_purchase_timestamp) AS month,
  222 p.payment_type,
  223 COUNT(*) AS num_orders
  224 FROM \underline{\text{`scaler-dsml-sql-gcp.target\_case\_study.orders'}} o
  225  JOIN `scaler-dsml-sql-gcp.target_case_study.payments` p
  ON o.order_id = p.order_id
  227 GROUP BY year, month, p.payment_type
  228 ORDER BY year, month, p.payment_type;
                                                                                           Query results
                                                                      ≛ SAVE RESULTS ▼
   JOB INFORMATION
                       RESULTS
                                    CHART
                                               JSON
                                                          EXECUTION DETAILS
                                                                                EXECUTION GRAPH
                                        payment_type 🔻
                                                                  num_orders ▼
 Row
                         month -
     1
                  2016
                                        credit_card
                                                                             3
     2
                  2016
                                    10
                                                                             63
    3
                  2016
                                    10
                                                                            254
                                        credit_card
     4
                  2016
                                                                             2
                                    10
     5
                  2016
                                        voucher
                                                                             23
     6
                  2016
                                    12
                                        credit_card
                                                                             1
                                    1
     7
                  2017
                                        UPI
                                                                            197
                                    1
    8
                  2017
                                        credit_card
                                                                            583
                                        debit_card
                                                                             9
    10
                  2017
                                        voucher
                                                                             61
```

# b) Find the no. of orders placed on the basis of the payment installments that have been paid.



#### 7) Actionable insights and recommendations:

- In conclusion, the analysis of the target dataset revealed several key insights and actionable recommendations to drive growth and enhance operational efficiency.
- The total & average value of order price for state SP is comparatively higher than the other states. And
  the difference between SP and RJ is 64%. So it is better to focus on regional purchases for overall
  growth.
- I would suggest Integrate augmented reality technology into the target (e-commerce) platform to offer immersive shopping experiences, allowing customers to visualise products in their real-world environment which will enhance the purchase rate.
- Also observed that most of the customers were ordered using credit cards. Providing extra offers for those who make payments using credit cards would definitely increase the number of purchases for ex: 20% off, up to \$ 350.
- Finally Customer satisfaction directly correlates with efficient delivery times and accurate order fulfilment.