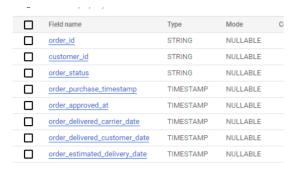
# Business Case: Target SQL

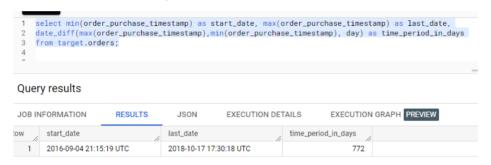
24 October 2022 19:27

### 1. Initial exploration of dataset like checking the characteristics of data

1. data type of columns in a table



2. Time period for which the data is given



3. Cities and States covered in the dataset

```
1 select distinct c.customer_city as cities,
2 c.customer_state as states from target.orders o join target.customers c
3 on o.customer_id = c.customer_id
4 order by cities , states;
```

## Query results

JOB INFORMATION RESULTS		JSON	EXECUTION DE	TAILS EXEC	
Row	cities	//	states	//	
1	abadia dos doura	ados	MG		
2	abadiania		GO		
3	abaete		MG		
4	abaetetuba		PA		
5	abaiara		CE		
6	abaira		BA		
7	abare		BA		
8	abatia		PR		
9	abdon batista		SC		
10	abelardo luz		SC		
11	abrantes		BA		
12	abre campo		MG		
13	abreu e lima		PE		
14	acaiaca		MG		
15	acailandia		MA		

#### 2. In-depth Exploration:

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

```
6 select count(order_id) as number_of_orders,
7 extract( month from order_purchase_timestamp) as order_month,
8 extract(year from order_purchase_timestamp) as order_year
9 from target.orders
10 group by order_year, order_month
11 order by order_year;
```

#### Query results

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS
Row	number_of	order_month	order_year	
1	324	10	2016	
2	4	9	2016	
3	1	12	2016	
4	7544	11	2017	
5	5673	12	2017	
6	2404	4	2017	
7	4026	7	2017	
8	4631	10	2017	
9	3245	6	2017	
10	4285	9	2017	
11	1780	2	2017	
12	800	1	2017	
13	4331	8	2017	
14	2682	3	2017	
15	3700	5	2017	

Exploring the resultant data in excel, we see that there is a increase trend in number of orders year on year basis.

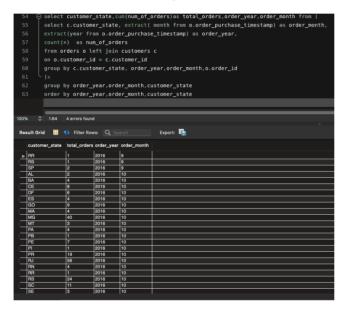
SUM of number_of_orders	order_year		
order_month	2016	2017	2018
1		800	7269
2		1780	6728
3		2682	7211
4		2404	6939
5		3700	6873
6		3245	6167
7		4026	6292
8		4331	6512
9	4	4285	16
10	324	4631	4
11		7544	
12	1	5673	
sum of orders	329	45101	54011

	^	D D	0	U
1	SUM of number_of_orders	order_year		
2	order_month	2016	2017	2018
3	1		800	7269
4	2		1780	6728
5	3		2682	7211
6	4		2404	6939
7	5		3700	6873
8	6		3245	6167
9	7		4026	6292
10	8		4331	6512
11	9	4	4285	16
12	10	324	4631	4
13	11		7544	
14	12	1	5673	
15	sum of orders	329	45101	54011
16				

Judging from the data available, we see highest peak of number of orders in month of October-November 2017 to January 2018 (towards er

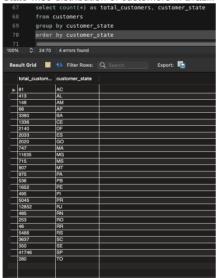
What time do Brazilian customers tend to buy (Dawn, Morning, Afternoon or Night)? Mostly in afternoon and followed by night, morning and least at dawn respectively

- 3. Evolution of E-commerce orders in the Brazil region:
  - 1. Get month on month orders by region, states



2. How are customers distributed in Brazil?

State wise distribution of customers in Brazil:



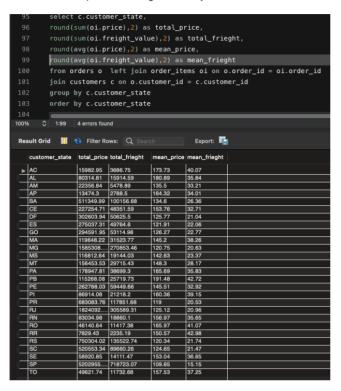
- 4. Impact on Economy: Analyze the money movemented by e-commerce by looking at order prices, freight and others.
  - 1. Get % increase in cost of orders from 2017 to 2018 (include months between Jan to Aug only)

The price of orders from 2017 to 2018 from Jan to Aug is 57.85

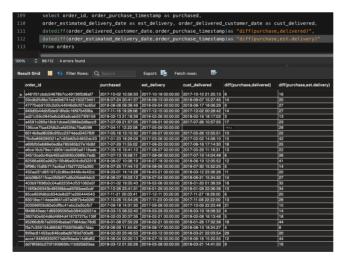
```
Heat total_price, order_year, orda_price_2017, round(((total_price - total_price_2017)/total_price));

lect round(sum|cl.price),2) as total_price, extract( year from o.order_purchase_timestamp) as total_price, extract( year from o.order_purchase_timestamp) as dotal_price, 2017
on orders to left join order_items of
o.order_ld=oil.order_id
Ⅲ ♦ Filter Rows: Q Search Export: 🖳
```

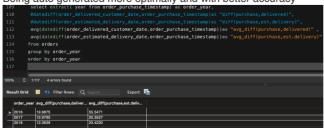
2. Mean & Sum of price and freight value by customer state



- 5. Analysis on sales, freight and delivery time
  - 1. Calculate days between purchasing, delivering and estimated delivery

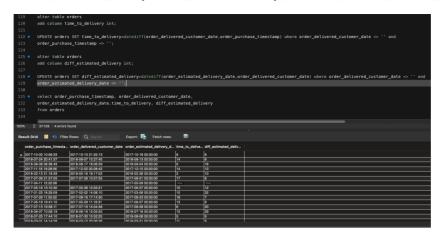


The average difference between the dates year on year decreased over time, the difference is more visible between 2016 and 20 Specially in difference between estimated delivery date and actual delivered date suggesting that the estimated delivery date is n Being auto generated more optimally and with better accuracy

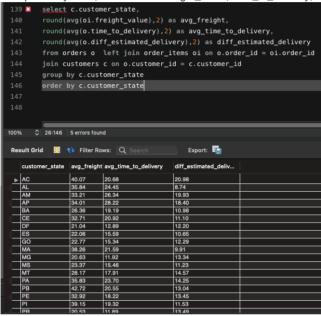


2. Create columns:

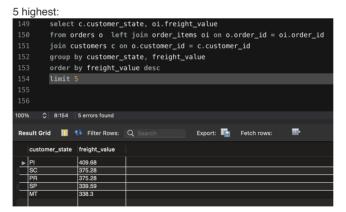
- time\_to\_delivery = order\_purchase\_timestamp-order\_delivered\_customer\_date
- diff\_estimated\_delivery = order\_estimated\_delivery\_date-order\_delivered\_customer\_date



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

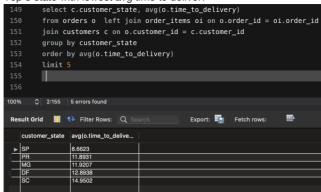


- 4. Sort the data to get the following:
  - 1. Top 5 states with highest/lowest average freight value sort in desc/asc limit 5

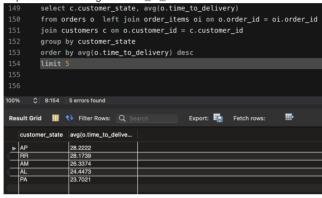


2. Top 5 states with highest/lowest average time to delivery

Top 5 state with lowest avg time to deliver:

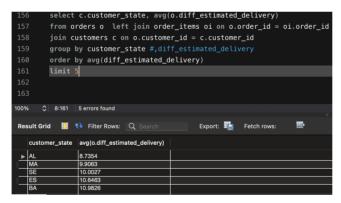


Top 5 states with highest time\_to\_deliver:

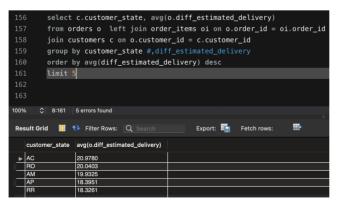


3. Top 5 states where delivery is really fast/ not so fast compared to estimated date

5 states with really fast delivery:(on an avg over the years)



5 states with not so fast delivery: (on an avg over the years)

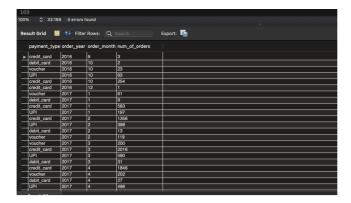


We can take a look at the logistics on the top 5 states with fastest delivery time on an average and can implement similar optimization c Portray the estimated delivery date as per actual delivery in the states with not so fast delivery.

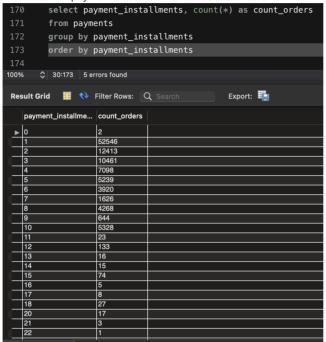
6. Payment type analysis:

1. Month over Month count of orders for different payment types

```
select p.payment_type, extract( year from order_purchase_timestamp) as order_year,
extract(month from o.order_purchase_timestamp) as order_month, count(*) as num_of_orders
from orders o left join payments p on o.order_id=p.order_id
where payment_type is not null
group by payment_type.order_year, order_month
order by order_year,order_month
```



2. Distribution of payment instalments and count of orders



#### Actionable Insights:

- The most orders comes around year end and new year. Target can scale up their infrastructure specially during this time and can sales even further
- Few states have considerable delta between estimated delivery date and actual delivery date. We can work with our logistics part network and reduce the time between order purchase and customer delivery. We can also tweak/optimize our estimated delivery delivery dates more accurately
- We see that there are many orders purchased via credit cards and we can partner with leading credit card provider in the country made via credit cards to even boost sales

#### Recommendations:

- To work with our delivery partners to reduce the time delta between purchase date and customer delivery date.
- For the states with highest freight charges, we can find alternate routes or work with freight institutions for exclusive partnerships time of delivery to customers
- Since there was a high YoY percentage of cost increase, we need to find ways to bring the costs down for a better sustainable bu on customer experience