PROJECT : SQL

BUSINESS CASE STUDY- THE OPERATIONS OF TARGET IN BRAZIL

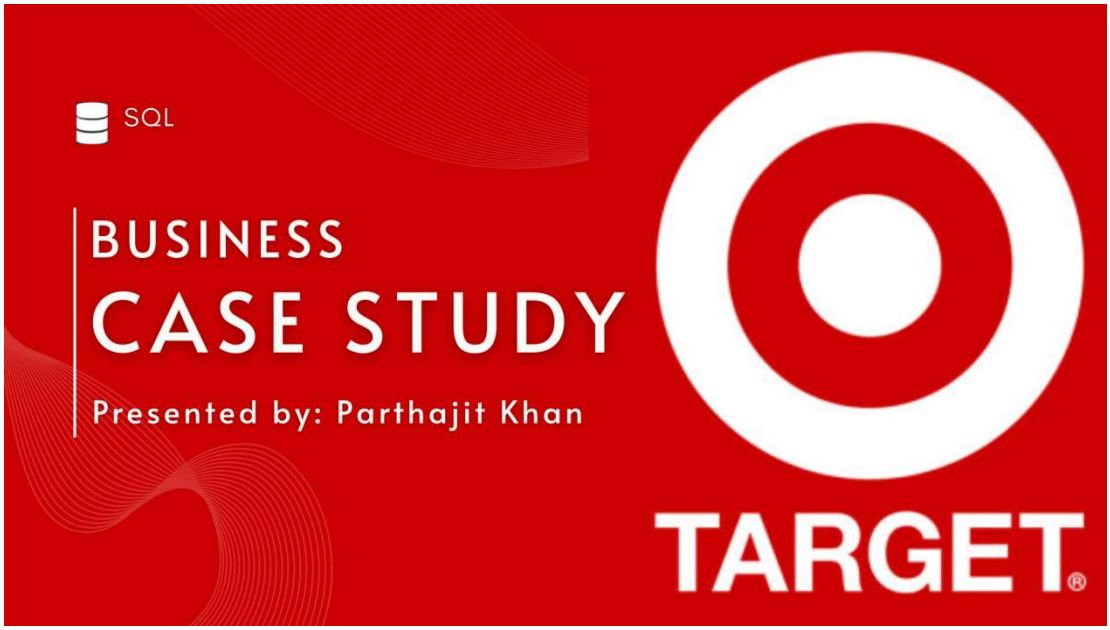


Table Of Contents :

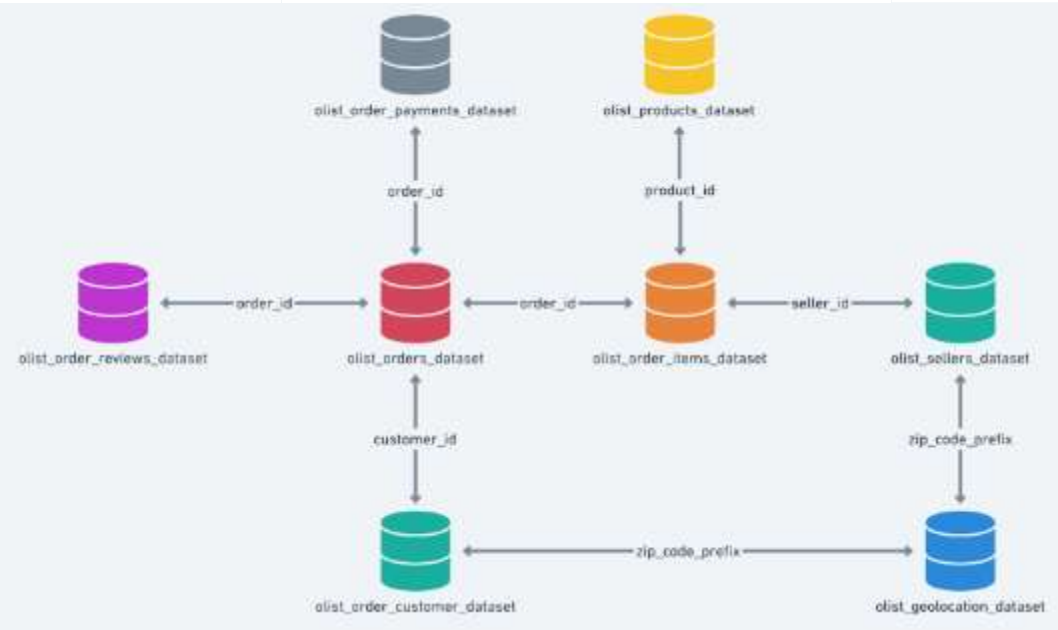
Company Description :

- 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. The dataset offers a comprehensive view of various dimensions including the order status, price, payment and freight performance, customer location, product attributes, and customer reviews.

- By analyzing this extensive dataset, it becomes possible to gain valuable insights into Target's operations in Brazil. The information can shed light on various aspects of the business, such as order processing, pricing strategies, payment and shipping efficiency, customer demographics, product characteristics, and customer satisfaction levels

Dataset :



QUESTIONS :-

1- Exploratory analysis :

a. Checking Data type of all columns in the "customers" table :

QUERY :

```
select column_name,data_type
 from data-science-409312.TARGET.INFORMATION_SCHEMA.COLUMNS
where table_name="Customer"
```

OUTPUT :

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	column_name	data_type
1	customer_id	STRING
2	customer_unique_id	STRING
3	customer_zip_code_prefix	INT64
4	customer_city	STRING
5	customer_state	STRING

INSIGHTS:

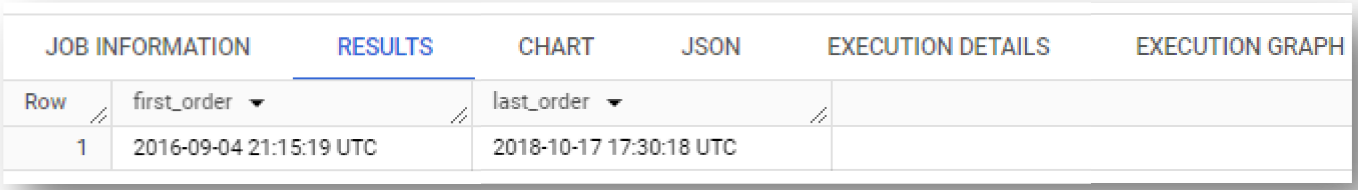
As we can see from the above schema of “customers” table there are 5 columns present & there datatypes are-

- **“customer_zip_code_prefix”-** It is the only column which contain **Integer** data type values out of 5 columns.
- **“customer_id”,”customer_unique_id”,”customer_city”,”customer_state”**- All these 4 columns are having same data type that is **String.**

b. Checking time range between which the orders were placed :

Query :

```
select min(order_purchase_timestamp) as first_order,
max(order_purchase_timestamp) as last_order
from `TARGET.Orders`
```



JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	first_order	last_order
1	2016-09-04 21:15:19 UTC	2018-10-17 17:30:18 UTC

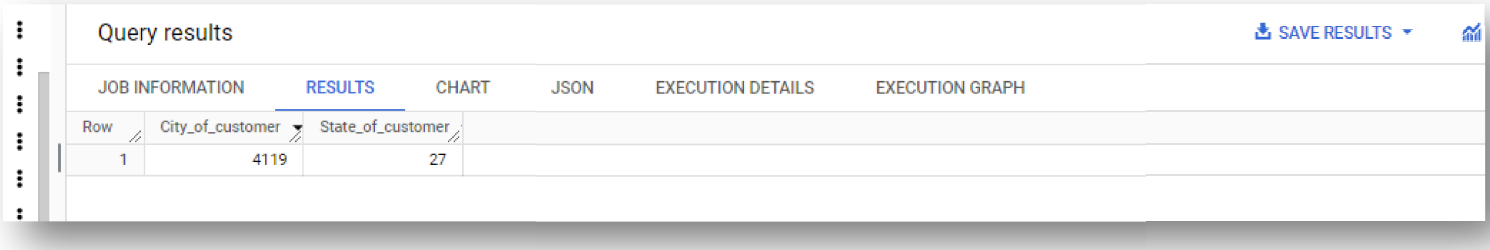
INSIGHTS:

- From the above query we can cleary say that the orders placed in the time range between 2016 & 2018 & also in our given context it has been clearly mentioned that the orders placed between 2016 & 2018.

c. Counting the Cities & States of customers who ordered during the given period :-

Query-

```
select count(distinct customer_city) as City_of_customer,
count(distinct customer_state) as State_of_customer
from `TARGET.Customer`
where customer_id in (
select customer_id from `TARGET.Orders`)
```



INSIGHTS:

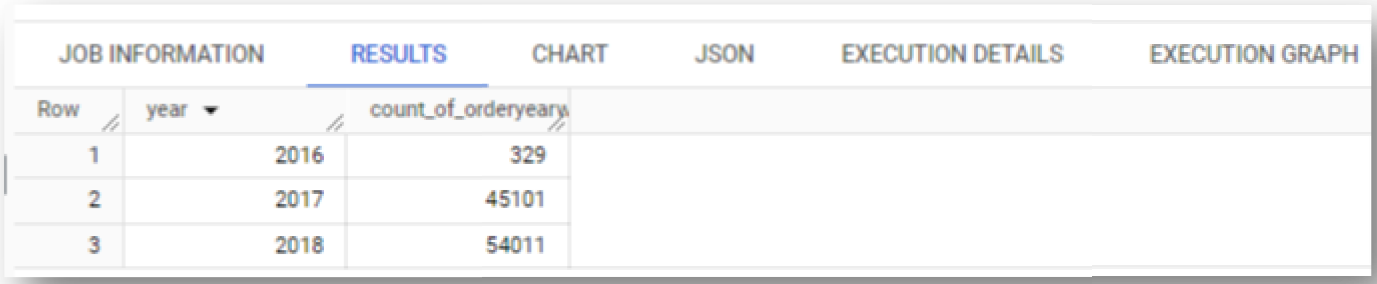
- From the above query we can observe that there are total **4119** nos of cities & **27** nos of states from where the Customers made the order during the time range of 2016 to 2018.

2. In-depth Exploration:

a. Checking is there a growing trend in the no. of orders placed over the past years? :

Query:

```
select extract(year from  order_purchase_timestamp) as year,
count(order_id) count_of_orderyearwise
from `TARGET.Orders`
 group by 1
 order by 1
```



JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	year	count_of_orderyearw
1	2016	329
2	2017	45101
3	2018	54011

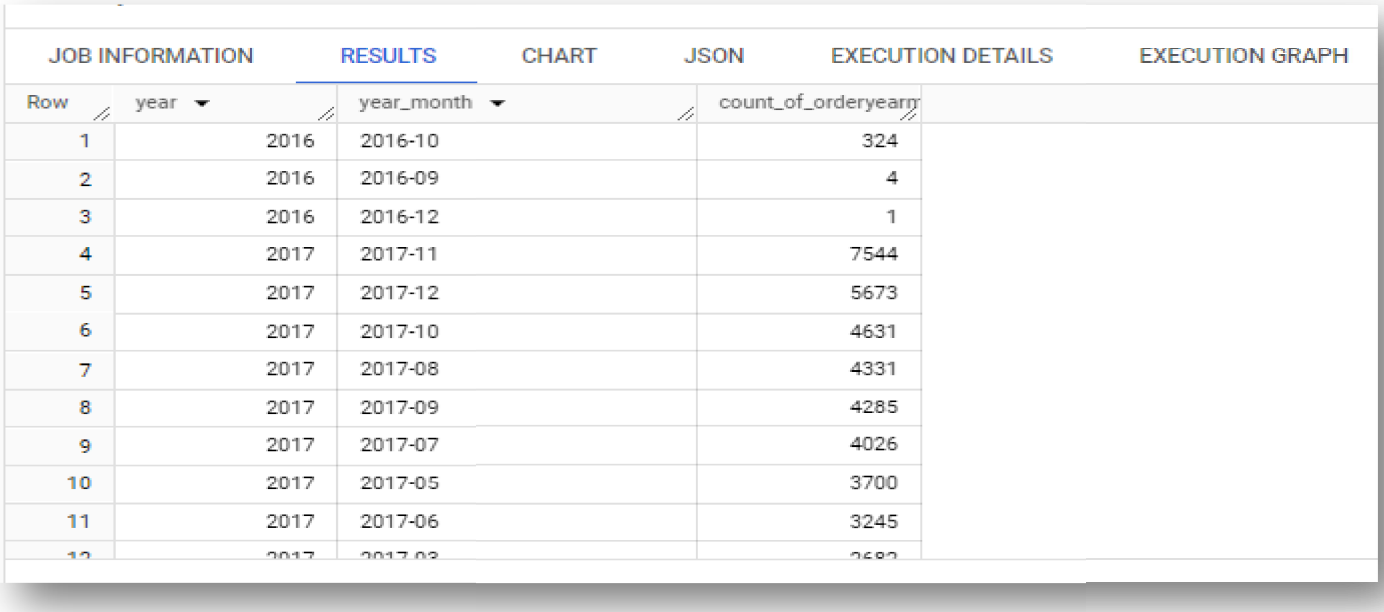
Inference:

- Here we can clearly observe that in there is a growing trend in terms of count of order placed year wise.

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

Query:

```
Select
Extract (year from order_purchase_timestamp) as year
format_datetime("%Y-%m",order_purchase_timestamp) as year_month,
count(order_id) count_of_orderyearmonthwise
from `TARGET.Orders`
 group by 1,2
 order by 1,3 desc
```



JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	year	year_month	count_of_orderyearm
1	2016	2016-10	324
2	2016	2016-09	4
3	2016	2016-12	1
4	2017	2017-11	7544
5	2017	2017-12	5673
6	2017	2017-10	4631
7	2017	2017-08	4331
8	2017	2017-09	4285
9	2017	2017-07	4026
10	2017	2017-05	3700
11	2017	2017-06	3245
12	2017	2017-03	2682

INSIGHTS:

We have seen that the count of order is increasing year wise, but when we focus year wise monthly count of order , we can notice

These following observations:

- In **2016-sep** the order count was extremely low –it may be because Target just started its operation in Brazil in 2016.so initially it may be due to lack of proper setup or customer recognition the sales were low.but if we focus on "**2017 –sep**" there are a significant amount of order can be seen. But "**2018-sep**" again it is low, because Target was closing its operation this year, so it may not taking any extra order.
- In 2017 **October-November-December**—we can observe progressive amount of order, It may due to the famous festivals in Brazil like Black Friday, Christmas, New Year's Eve which drive significant retail activity in Brazil.
- Overall it demonstrates a consistent & significant upward trend in the number of orders which indicating an increasing customer demand of TARGET's products in Brazil.

c. 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

Query-

```
with cte as (
select order_id,order_purchase_timestamp,extract(year from order_purchase_timestamp) as
year,cast(format_timestamp("%H",timestamp(order_purchase_timestamp)) as int64) as time
 from `TARGET.Orders`),
 cte2 as (
 select *,case when time>=0 and time <=6
        then "Dawn"
        when time >=7 and time <=12
        then "Mornings"
        when time>=13 and time <=18
        then "Afternoon"
        when time>=19 and time <=23
        then "Night"
        end as time_label
        from cte)

select time_label,count(order_id) as order_count from cte2
group by 1
order by 2 desc
```

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	time_label	order_count
1	Afternoon	38135
2	Night	28331
3	Mornings	27733
4	Dawn	5242

INSIGHTS :

- Based on the output , Brazilian customers tend to place the highest no of orders during the **Afternoon** hours followed by the **Night** hours.
- **Morning** hour also show a substantial volume of orders ,while **Dawn** hours have fewer orders placed as it is the sleeping time.
- This suggests that customers of Brazil are actively engaged in shopping during **Afternoon** , **Night** & **Morning.**

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

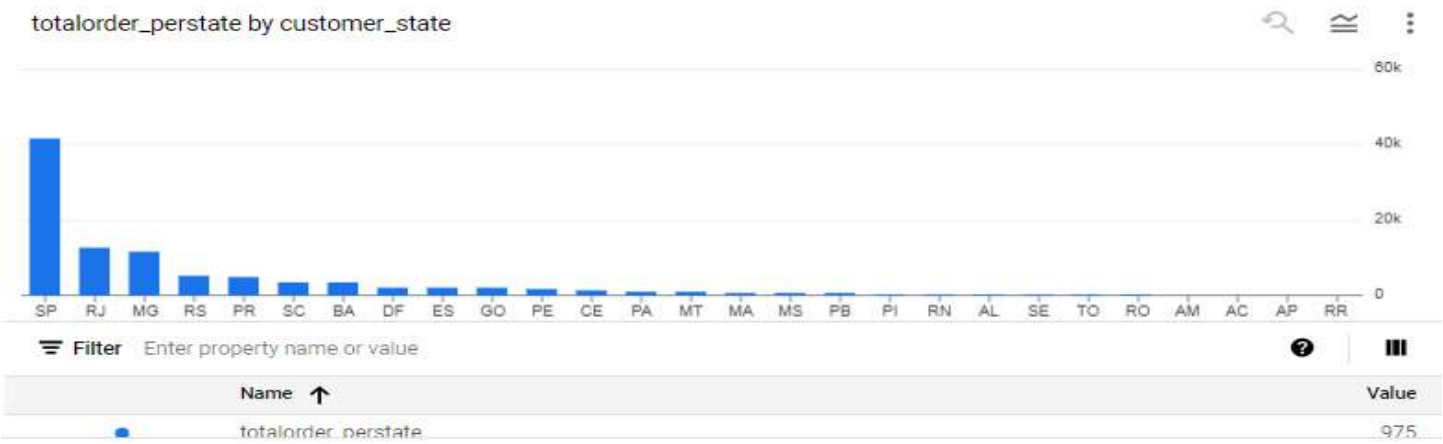
a. Checking the month on month no. of orders placed in each state

```
with cte as (
select format_datetime("%b",a.order_purchase_timestamp) as mnth,b.customer_state,a.order_id
from `TARGET.Orders` a
left join `TARGET.Customer` b
on a.customer_id=b.customer_id),
cte2 as (select customer_state,mnth,count(order_id) as countoforder
 from cte
 group by 1,2
 order by 3 desc)

 select *,sum(countoforder)over(partition by customer_state) as totalorder_perstate
 from cte2
 order by 4 desc
```

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	customer_state	mnth	countoforder	totalorder_perstate
1	SP	Dec	2357	41746
2	SP	Jul	4381	41746
3	SP	May	4632	41746
4	SP	Oct	1908	41746
5	SP	Jan	3351	41746
6	SP	Feb	3357	41746
7	SP	Apr	3967	41746
8	SP	Jun	4104	41746
9	SP	Mar	4047	41746
10	SP	Nov	3012	41746
11	SP	Aug	4982	41746
12	SP	Sep	1648	41746



INSIGHTS:

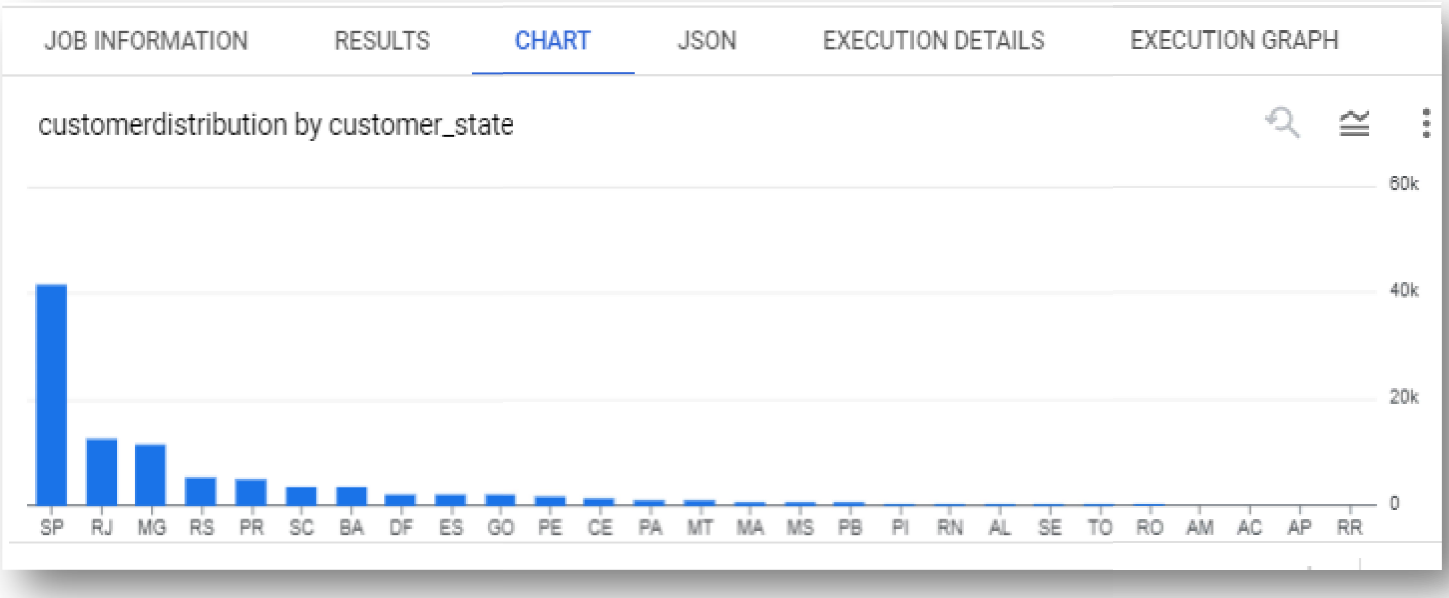
- In the above result , it has been shown the numbers of order placed in each state from highest to lowest for each state.
  “SP” is the state where the highest amount of order has been placed.

Q1. How are the customers distributed across all the states:

```
select customer_state,
count(customer_id) as customerdistribution
from `TARGET.Customer`
GROUP BY 1
order by 2 desc
```

JOB INFORMATION | RESULTS | CHART | JSON

Row	customer_state	customerdistribution
1	SP	41746
2	RJ	12852
3	MG	11635
4	RS	5466
5	PR	5045
6	SC	3637
7	BA	3380
8	DF	2140
9	ES	2033
10	GO	2020
11	PE	1652
12	CE	1336



INSIGHTS:

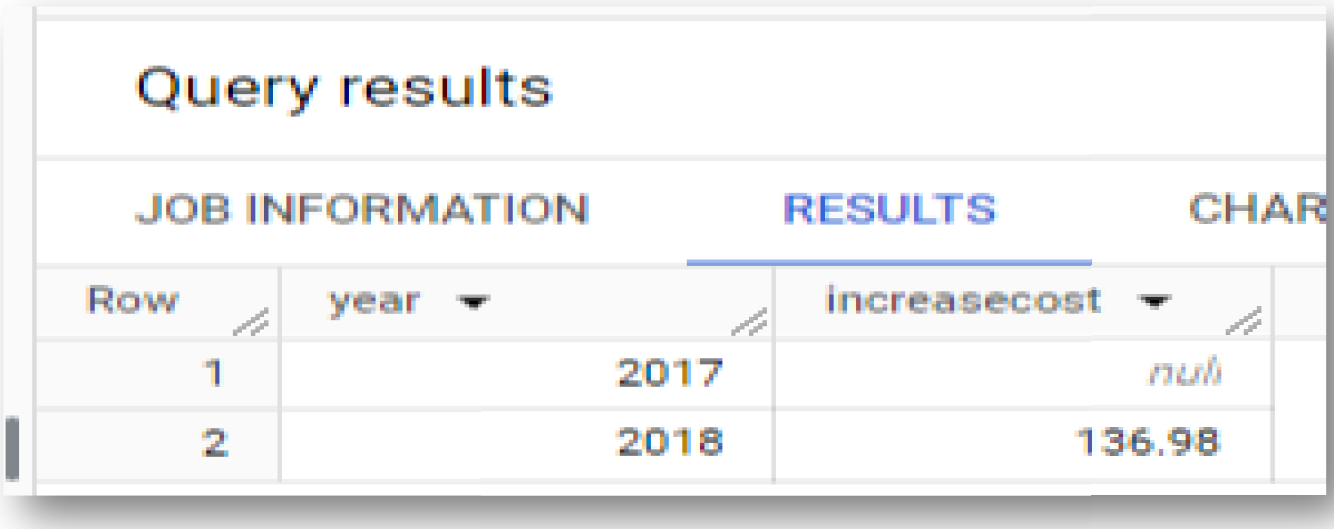
- From above result it can be seen how the customers are distributed over states. E.g.-“SP” state has the highest numbers of customers etc & so on --------This analysis will be helpful in decision making of where to focus more & more.

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

Q1. Get the % increase in the cost of orders from year 2017 to 2018 (including months between Jan to Aug only)

```
 with cte as (
select a.year,sum(payment_value) as ordercost
from (select order_id ,year,mnth from
(select order_id,cast(extract(year from order_purchase_timestamp) as int64) as year,extract(month from
order_purchase_timestamp) as m,format_datetime("%b",order_purchase_timestamp) as mnth
from `TARGET.Orders`) v
where  year >=2017 and year<=2018 and m>=1 and m<=8) as a
join `TARGET.Payments`  as b
on a.order_id=b.order_id
group by 1
order by 1)

select year,round(((ordercost-lag(ordercost,1)over(order by ordercost))
            /(lag(ordercost,1)over(order by ordercost)))*100,2) as increasecost
            from cte
```



INSIGHTS:

- From the above result we notice that there are 136.98% approx **137**% order cost increased from the year “2017” to “2018” which indicating that Significant Growth in Demand or Sales Volume,market expansion, improved business performance.

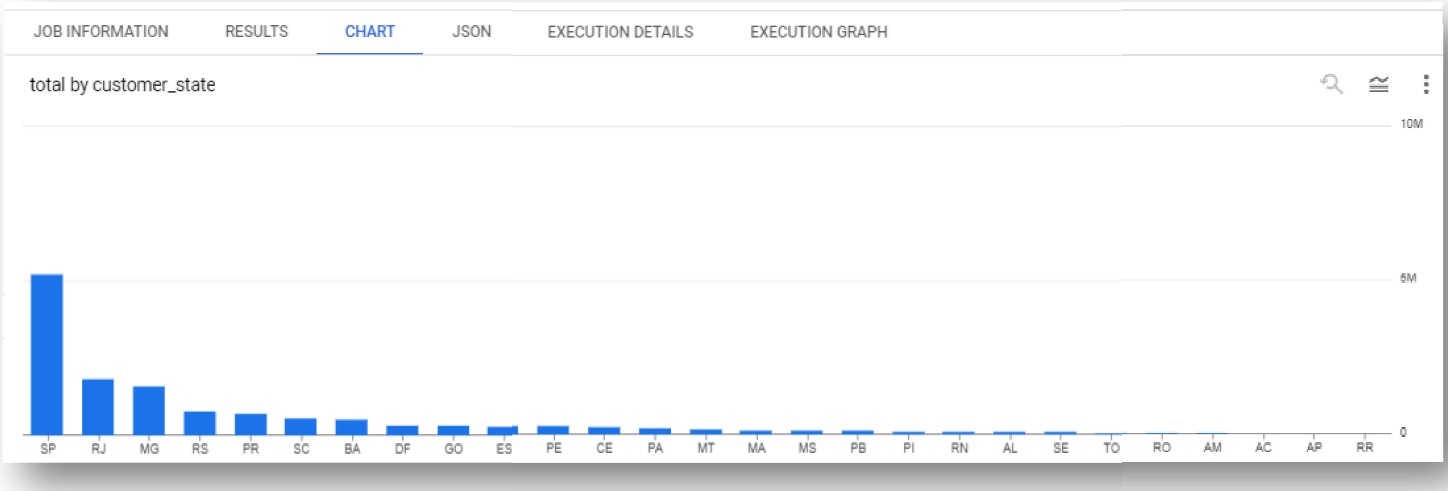
Q2. Calculate the Total & Average value of order price for each state

Query1(calculating total):

```
select c.customer_state,round(sum(price),2) as total
from `TARGET.Customer` c
join `TARGET.Orders` o
on c.customer_id=o.customer_id
join `TARGET.Order_item` ot
on o.order_id=ot.order_id
group by 1
order by 2 desc
```

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	customer_state	total
1	SP	5202955.05
2	RJ	1824092.67
3	MG	1585308.03
4	RS	750304.02
5	PR	683083.76
6	SC	520553.34
7	BA	511349.99
8	DF	302603.94
9	GO	294591.95
10	ES	275037.31
11	PE	262788.03

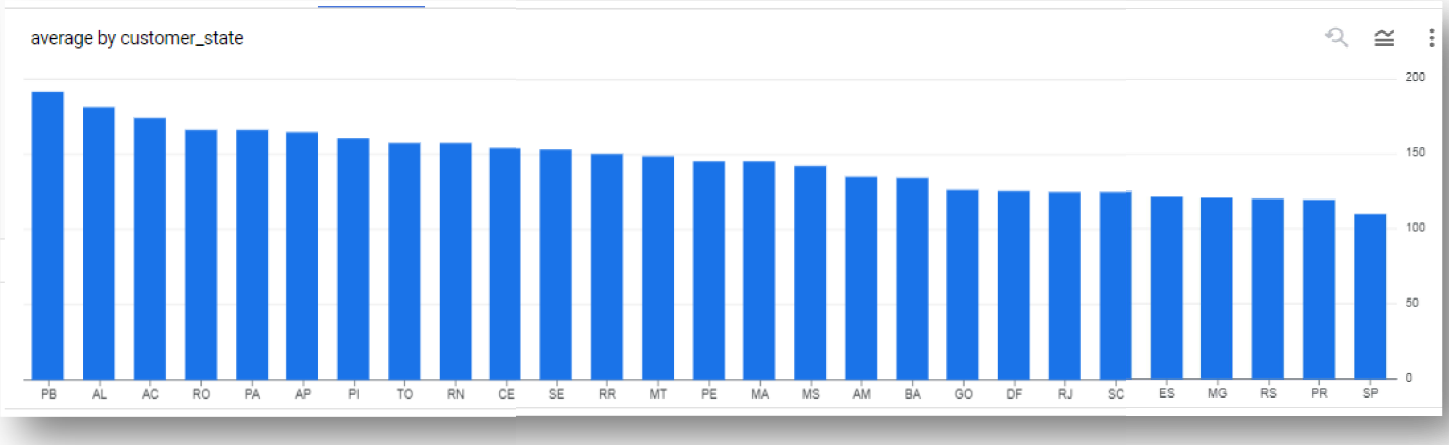


Query 2((calculating average):

```
select c.customer_state,round(avg(price),2) as average
from `TARGET.Customer` c
join `TARGET.Orders` o
on c.customer_id=o.customer_id
join `TARGET.Order_item` ot
on o.order_id=ot.order_id
group by 1
order by 2 desc
```

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	customer_state	average
1	PB	191.48
2	AL	180.89
3	AC	173.73
4	RO	165.97
5	PA	165.69
6	AP	164.32
7	PI	160.36
8	TO	157.53
9	RN	156.97
10	CE	153.76
11	SE	153.04



INSIGHTS:

- **State with the Highest Total Order Price :**
  - This states(e.g-"**SP**" likely has a large customer base or a higher volume of orders. The total order price being the highest suggests that the overall demand in this state is significant.

- It might indicate that the company has a strong market presence in these states, with many repeat customers or high-frequency orders

- **State with the Highest Average Order Price :**
  - This states(e.g-"**PB**" likely sees fewer orders, but each order is of higher value. This could indicate that customers in this state tend to purchase more expensive items or make bulk purchases.
  - The high average order price might also suggest a wealthier customer base or a market where premium products or services are more popular
- The company may need to employ different marketing and sales strategies in these states. For the state with the highest total order price, the focus could be on maximizing volume, while for the state with the highest average order price, the focus could be on promoting premium offerings or upselling

  ---------- These inferences can guide strategic decisions, such as where to focus marketing efforts, how to tailor products or services to meet regional demands, and how to allocate resources effectively to maximize both revenue and profit.

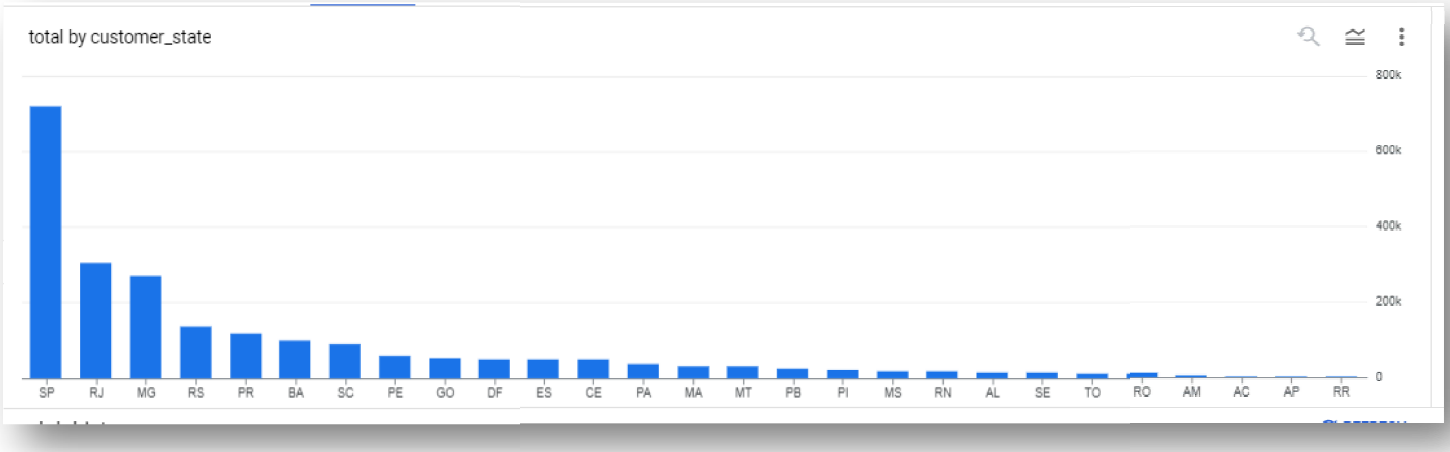
c. Calculate the Total & Average value of order freight for each state :

QUERY1(calculating total):

```
select c.customer_state,
round(sum(freight_value),2) as total
from `TARGET.Customer` c
join `TARGET.Orders` o
on c.customer_id=o.customer_id
join `TARGET.Order_item` ot
on o.order_id=ot.order_id
group by 1
order by 1
```

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	customer_state	total
1	SP	718723.07
2	RJ	305589.31
3	MG	270853.46
4	RS	135522.74
5	PR	117851.68
6	BA	100156.68
7	SC	89660.26
8	PE	59449.66
9	GO	53114.98
10	DF	50625.5
11	ES	49764.6

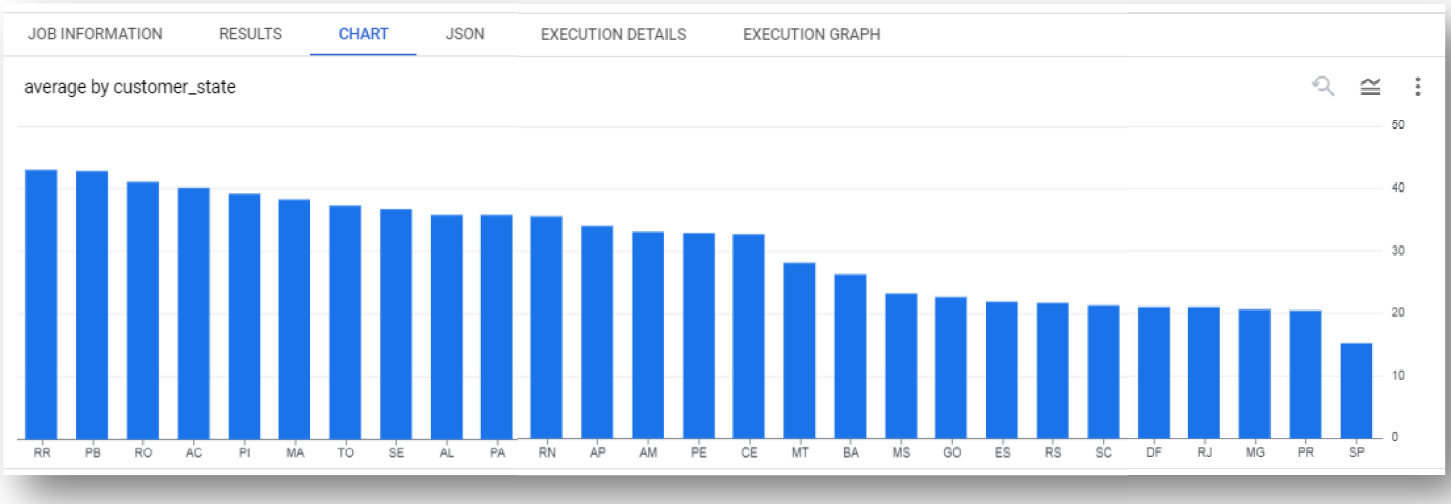


QUERY2(calculating average):

```
select c.customer_state,
round(avg(freight_value),2) as average
from `TARGET.Customer` c
join `TARGET.Orders` o
on c.customer_id=o.customer_id
join `TARGET.Order_item` ot
on o.order_id=ot.order_id
group by 1
order by 1
```

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	customer_state	average
1	RR	42.98
2	PB	42.72
3	RO	41.07
4	AC	40.07
5	PI	39.15
6	MA	38.26
7	TO	37.25
8	SE	36.65
9	AL	35.84
10	PA	35.83
11	RN	35.65
12	AP	34.01



INSIGHTS:

- **Operational Focus**: The state(“e.g-**“SP”)** with the highest total freight value is operationally focused on high volume, suggesting it might benefit from further investment in logistics infrastructure to enhance efficiency and reduce costs
- **Customer/Product segmentation**: The states(e.g-“**RR”**) with the highest average freight value likely deals with customers or products that need special shipping, like longer distances, heavy items, or fragile goods. This means the focus is on delivering quality or meeting specific needs, rather than just shipping a lot of items cheaply. This leads to higher costs per shipment, reflecting a different strategy focused on specialized services
- **Freight Cost Management**: Identifying areas where freight costs can be optimized, such as consolidating shipments or negotiating better rates with carriers, especially in the state with high total freight value.
- **Customer Pricing Strategies**: Adjusting pricing strategies to reflect the higher freight costs in the state with high average freight value(e.g-“**RR”)**, possibly through tiered shipping options or incentives for choosing more economical shipping methods

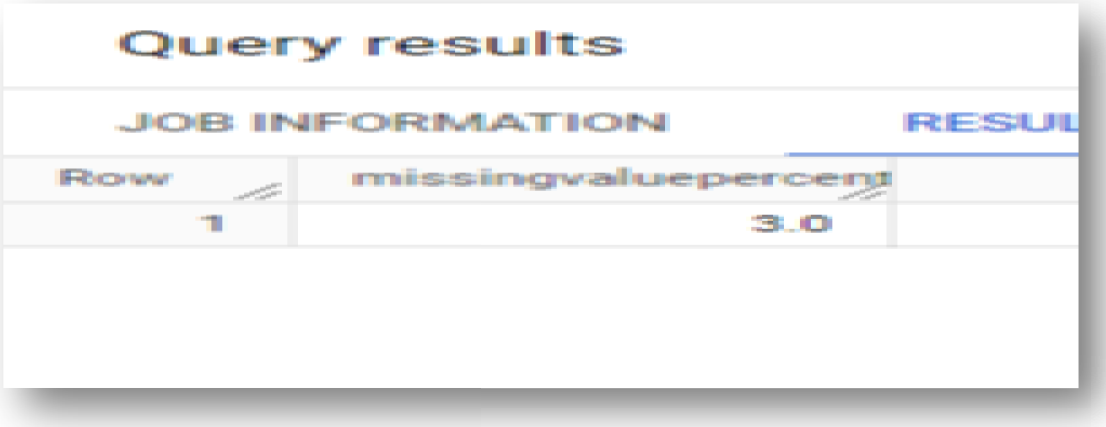
5. Analysis based on sales, freight and delivery time:

a. Finding the no. of days taken to deliver each order from the order’s purchase date as delivery time & finding the difference (in days) between the estimated & actual delivery date of an order

Query:

Checking Null values in Orders table-

```
select
round((count(order_id)/(select count(order_id) from `TARGET.Orders` ))*100) as missingvaluepercentage
from `TARGET.Orders`
 where order_delivered_customer_date  is null
```



Query results

JOB INFORMATION | RESUL

Row	missingvaluepercent
1	3.0

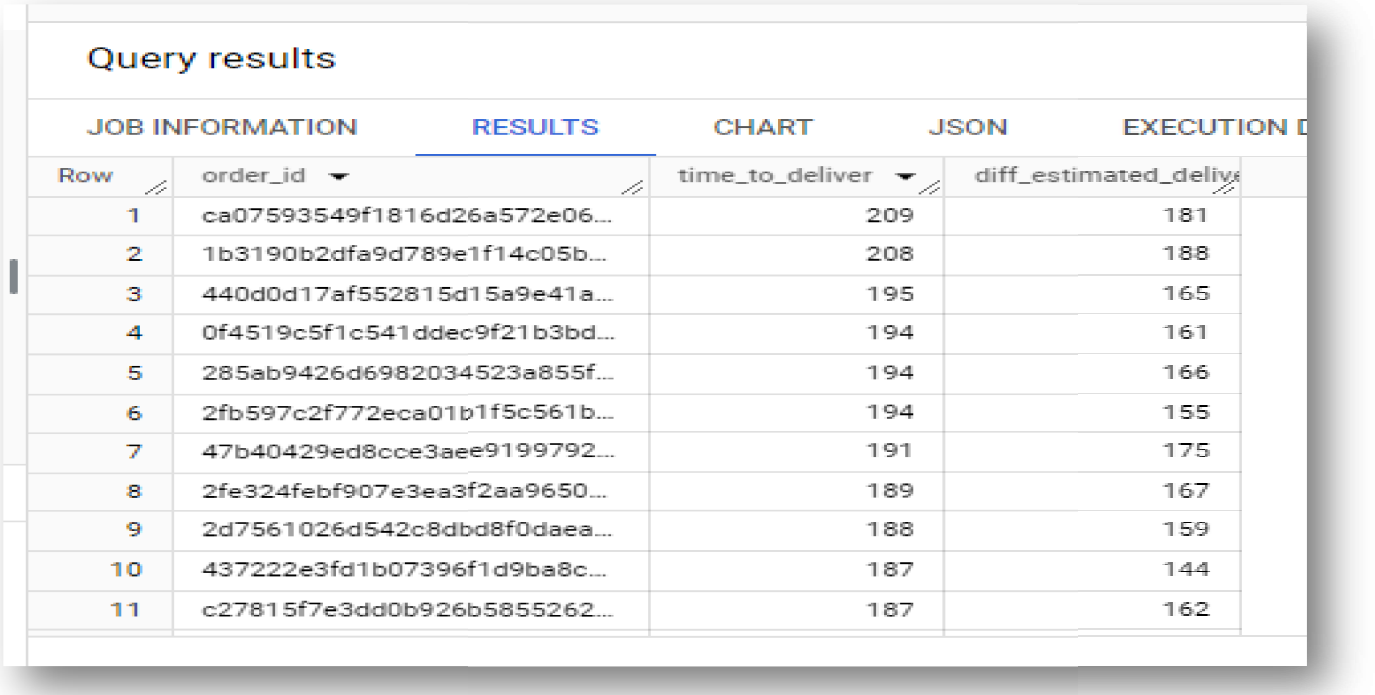
There are only **3%** missing values present in "**order_delivered_customer_date"** column , so it can be ignored.

Main query:

- **time_to_deliver** = order_delivered_customer_date - order_purchase_timestamp
- **diff_estimated_delivery** = order_delivered_customer_date - order_estimated_delivery_date

```
with cte as (select * from `TARGET.Orders`
where order_delivered_customer_date is not null)

select order_id,date_diff(order_delivered_customer_date,order_purchase_timestamp,day) as time_to_deliver,
abs(date_diff(order_delivered_customer_date,order_estimated_delivery_date,day)) as diff_estimated_delivery
 from cte
order by 2 desc
```



Query results

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION D

Row	order_id	time_to_deliver	diff_estimated_deliv
1	ca07593549f1816d26a572e06...	209	181
2	1b3190b2dfa9d789e1f14c05b...	208	188
3	440d0d17af552815d15a9e41a...	195	165
4	0f4519c5f1c541ddec9f21b3bd...	194	161
5	285ab9426d6982034523a855f...	194	166
6	2fb597c2f772eca01b1f5c561b...	194	155
7	47b40429ed8cce3aee9199792...	191	175
8	2fe324febf907e3ea3f2aa9650...	189	167
9	2d7561026d542c8dbd8f0daea...	188	159
10	437222e3fd1b07396f1d9ba8c...	187	144
11	c27815f7e3dd0b926b5855262...	187	162

INSIGHTS: From the upper result we can observe that it has taken maximum 209 days to deliver a order from the purchase date(by arranging in descending order of time_to_deliver column) & maximum 188 more days to deliver an order from the estimated delivery date (diff_estimated_delivery)----from this we can infer that :

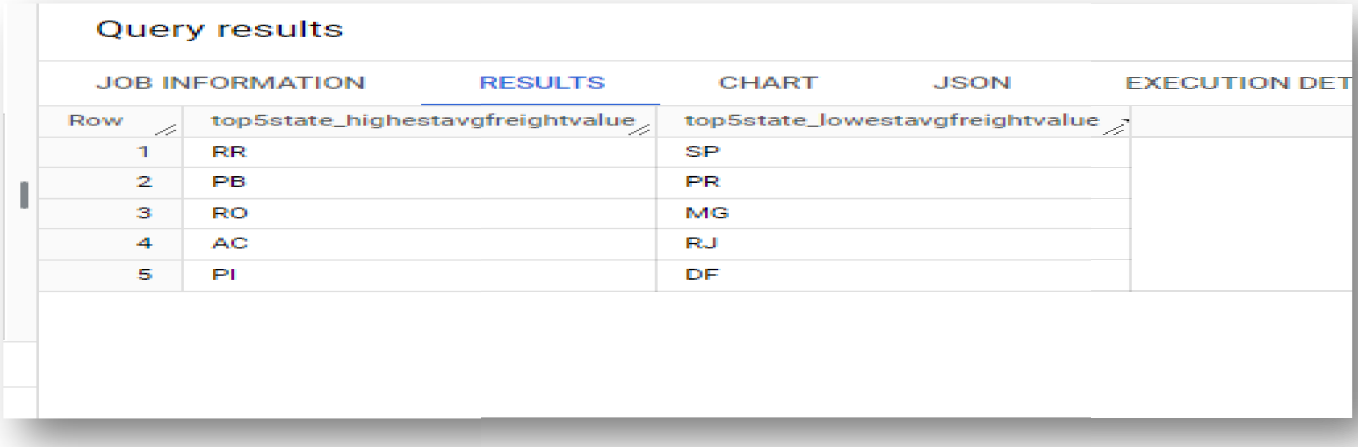
- **Delivery Efficiency**: If the actual delivery time is consistently longer than the estimated time, it indicates inefficiencies in the logistics or supply chain process.
- **Customer Satisfaction**: A significant difference between estimated and actual delivery dates can lead to customer dissatisfaction, as customers rely on accurate delivery estimates.
- **Performance of Carriers**: By comparing estimated and actual delivery times, we can assess the performance of shipping carriers. Frequent delays may suggest the need for better carrier partnerships or renegotiated terms.
- **Accuracy of Delivery Estimates**: If the difference is often large, it may suggest that the estimated delivery times are unrealistic or inaccurate, requiring a review of the estimation process.
- **Seasonal or External Influences**: Patterns in delays can reveal how external factors, such as holidays, weather conditions, or peak seasons, affect delivery times, allowing better planning during such periods.
- **Areas for Improvement**: Identifying common reasons for delays helps target specific areas for process improvements, like warehouse operations, packaging, or route optimization.
- **Benchmarking and Goal Setting**: The data can be used to set benchmarks and goals for delivery times, driving operational improvements and setting customer expectations more accurately.
  ----------These insights help in optimizing the delivery process, improving customer satisfaction, and refining logistics strategies for better performance

b. Finding out the top 5 states with the highest & lowest average freight value :

Query:

```
with cte as (
select c.customer_state,round(avg(freight_value),2) as average
from `TARGET.Customer` c
join `TARGET.Orders` o
on c.customer_id=o.customer_id
join `TARGET.Order_item` ot
on o.order_id=ot.order_id
group by 1)

select top5state_highestavgfreightvalue,top5state_lowestavgfreightvalue
from
(select customer_state as top5state_highestavgfreightvalue,
  row_number()over(order by average desc) as r
   from cte
order by average desc
limit 5) a
join
(select customer_state as top5state_lowestavgfreightvalue,
  row_number()over(order by average) as r
  from cte
order by average
limit 5) b
on a.r=b.r
```



Query results

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DET

Row	top5state_highestavgfreightvalue	top5state_lowestavgfreightvalue
1	RR	SP
2	PB	PR
3	RO	MG
4	AC	RJ
5	PI	DF

INSIGHTS:

From Top 5 States with Highest Average Freight Value:

- **Special Needs**: These states may require specialized shipping (e.g., heavy or bulky items, long distances).
  - **Action**: Invest in tailored logistics solutions or adjust pricing to cover higher costs.
- **Premium Services**: Customers might prefer faster or premium shipping options.
  - **Action**: Enhance or promote premium shipping services.
- **Cost Efficiency**: There may be opportunities to improve cost efficiency in logistics.
  - **Action**: Explore cost-saving measures or negotiate better rates with carriers.

From Top 5 States with Lowest Average Freight Value:

- **Economies of Scale**: Lower freight costs might indicate high shipping volumes.
  - **Action**: Optimize logistics to leverage scale for further cost reductions.
- **Operational Efficiency**: Efficient logistics operations may be contributing to lower costs.
  - **Action**: Maintain or enhance current logistics practices to ensure continued cost-effectiveness

c. Finding the top 5 states with the highest & lowest average delivery time :

Query:

- **delivertime** = order_delivered_customer_date - order_purchase_timestamp

```
with cte as (
select customer_state, avg(delivery_time) as avg_deliverytime
from (select
    customer_id,
    date_diff(order_delivered_customer_date,order_purchase_timestamp,day) as delivery_time
 from `TARGET.Orders`
 where order_delivered_customer_date is not null
) o
join `TARGET.Customer` c
on o.customer_id=c.customer_id
group by 1)

select top5state_highestavg,top5state_lowestavg from
(select customer_state as top5state_highestavg,
row_number()over(order by avg_deliverytime desc) as r_no
```

```
from cte
order by avg_deliverytime desc
limit 5) as a
join
(select customer_state as top5state_lowestavg,
row_number()over(order by avg_deliverytime) as r_no
from cte
order by avg_deliverytime
limit 5) as b
on a.r_no=b.r_no
```

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	top5state_highestavg	top5state_lowestavg
1	RR	SP
2	AP	PR
3	AM	MG
4	AL	DF
5	PA	SC

INSIGHTS :

From Top 5 States with Highest Average Delivery Time:

- **Logistics Bottlenecks**: These states may have issues like poor infrastructure, remote locations, or inefficiencies in the supply chain.
  - **Inference**: There are delays or challenges in the delivery process that need addressing.
- **Carrier Performance Issues**: Certain carriers might be underperforming or not well-suited for these regions.
  - **Inference**: Carrier selection or performance needs review and optimization.
- **Geographical Challenges**: Longer distances could contribute to longer delivery times.
  - **Inference**: Geography plays a significant role in delivery delays.
- **Seasonal or External Factors**: High delivery times could be influenced by external factors such as bad weather, holidays, or peak seasons.
  - **Inference**: Delivery times are affected by predictable external conditions.

From Top 5 States with Lowest Average Delivery Time:

- **Efficient Logistics Operations**: These states likely have streamlined logistics and efficient processes.
  - **Inference**: Existing logistics strategies are effective and well-optimized.
- **Proximity to Distribution Centers**: Closer proximity to warehouses or distribution centers reduces delivery times.
  - **Inference**: Strategic location benefits delivery speed.
- **Reliable Carriers**: High performance of carriers or delivery partners in these areas.
  - **Inference**: Carrier selection is effective and reliable.
- **Good Infrastructure**: Well-developed transportation and infrastructure networks contribute to faster deliveries.
  - **Inference**: Infrastructure quality positively impacts delivery speed.

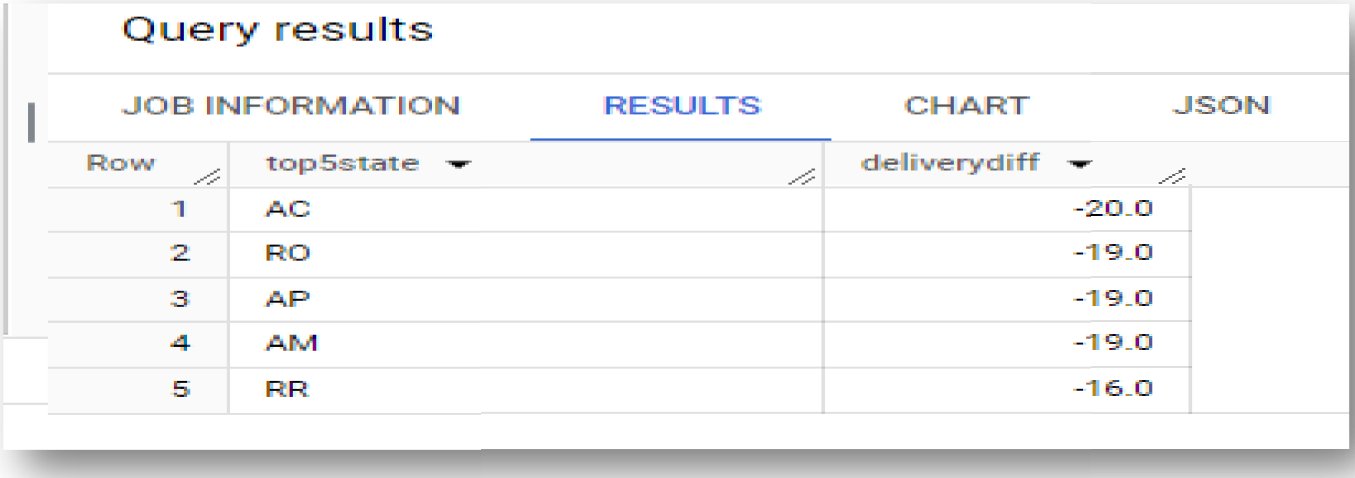
d. Finding the top 5 states where the order delivery is really fast as compared to the estimated date of delivery.

Query:

- **diff_estimated_delivery** = order_delivered_customer_date - order_estimated_delivery_date

```
with cte as (
select customer_state, avg(diff_estimated_delivery) as average_deliverydiff
from (select order_id,
    customer_id,
    (date_diff(order_delivered_customer_date,order_estimated_delivery_date,day)) as diff_estimated_delivery
 from `TARGET.Orders`
 where order_delivered_customer_date is not null
) a
join `TARGET.Customer` b
on a.customer_id=b.customer_id
group by 1)

select customer_state as top5state,round(average_deliverydiff) as deliverydiff
from cte
order by average_deliverydiff
limit 5
```



Row	top5state	deliverydiff
1	AC	-20.0
2	RO	-19.0
3	AP	-19.0
4	AM	-19.0
5	RR	-16.0

INSIGHTS: From the upper analysis based on sales ,freight value & delivery time we can observe some pattern in the output:

1. States with High Freight Values and Fast Delivery (e.g., RR,RO,AC etc)

- **Inference**: High freight costs are likely due to expedited or premium shipping services that ensure faster delivery than estimated dates. These states invest in speed, which drives up costs.

- **Operational Strategy**: For similar states, consider optimizing shipping costs by negotiating better rates for expedited services or exploring alternative carriers that offer competitive pricing without compromising speed.

2. States with High Delivery Time but Fast Compared to Estimates (e.g., AP, AM,etc)

- **Inference**: These states face inherent logistical or geographical challenges leading to high delivery times, but efficient in-transit operations allow them to deliver faster than estimated dates. This suggests a strong capability in managing expectations and last-mile delivery.
- **Operational Strategy**: For other states with similar challenges, improve initial logistics processes or adjust delivery estimates to be more realistic, enhancing overall delivery performance and customer satisfaction.

3. States with High Freight Values Without Fast Delivery

- **Inference**: High freight costs are driven by specialized logistics needs (e.g., heavy items, long distances) rather than speed. These states do not necessarily benefit from fast delivery despite high costs.
- **Operational Strategy**: For these states, focus on optimizing logistics to reduce costs, such as by consolidating shipments, adjusting packaging, or renegotiating terms with carriers for better bulk rates.

4. States with Fast Delivery and Moderate Costs

- **Inference**: These states achieve fast delivery relative to estimates without incurring high freight costs, indicating efficient logistics operations and cost-effective carrier partnerships.
- **Operational Strategy**: Emulate these states' practices in other regions, focusing on route optimization, carrier performance monitoring, and maintaining efficient order processing to keep delivery times low without increasing costs.0

5. States with low delivery time & low freight value

- These states indicates cost efficiency, enhance logistics, prioritize marketing, optimize supply chains, gain competitive advantage, manage inventory effectively, improve customer satisfaction, evaluate high-performing partners, and support sustainability initiatives

6. Analysis based on the payments:

a. Finding the month on month orders placed using different payment types :

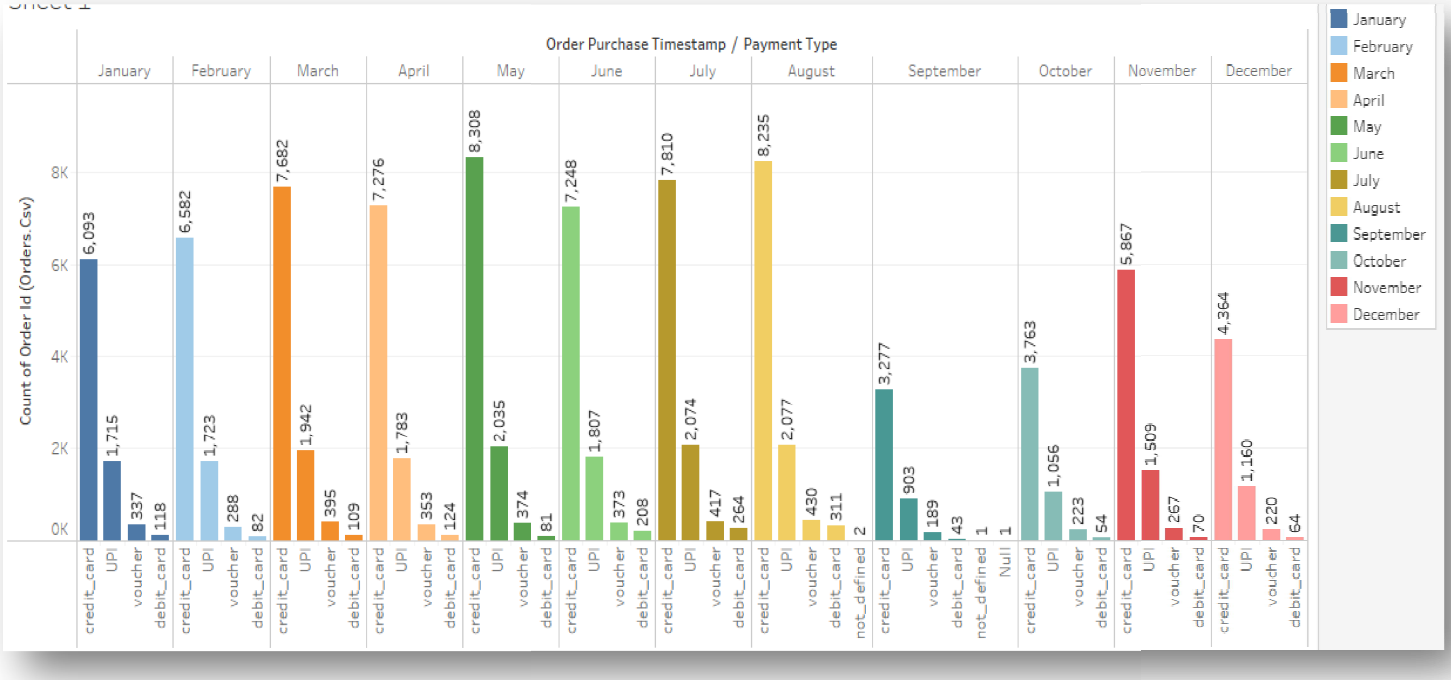
Query:

```
select mnth,payment_type,count_of_order
from(select format_timestamp("%b",order_purchase_timestamp) as
mnth,format_date("%m",order_purchase_timestamp),payment_type,count(a.order_id) as count_of_order
from `TARGET.Orders` a
join `TARGET.Payments` b
on a.order_id=b.order_id
group by 1,2,3
order by 2,4 desc)
```

Query results

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS

Row	mnth	payment_type	count_of_order
1	Jan	credit_card	6103
2	Jan	UPI	1715
3	Jan	voucher	477
4	Jan	debit_card	118
5	Feb	credit_card	6609
6	Feb	UPI	1723
7	Feb	voucher	424
8	Feb	debit_card	82
9	Mar	credit_card	7707
10	Mar	UPI	1942
11	Mar	voucher	591



INSIGHTS: After analysis the output & chart we can observe the following :-

- **Customer Payment Preferences:**

  - **Credit cards** are the most popular payment method, indicating that customers prefer or are more comfortable using them. This could suggest that offering promotions or incentives for credit card payments might be beneficial.

- **Payment Method Trends:**

  - **UPI** is the second most popular method, showing growing adoption. Businesses might consider enhancing UPI payment options or offering UPI-exclusive discounts to attract more users.
  - **Voucher** usage, while lower, still represents a segment of customers, suggesting that voucher promotions could drive additional sales or attract price-sensitive buyers.

- **Debit cards**, being the least used, may indicate a need for improved security features or convenience enhancements for debit card transactions.

- **Marketing and Promotions:**

  - Tailor marketing strategies to emphasize credit card and UPI payments, potentially offering rewards or discounts to encourage their use.
  - Consider introducing or enhancing voucher-based promotions to increase engagement with that payment method.

- **Customer Segmentation:**

  - Different payment methods might correlate with different customer segments. Understanding these segments can help in designing targeted marketing campaigns or customer loyalty programs.

- **Payment System Optimization:**

  - Ensure that payment systems for the most used methods (credit cards and UPI) are optimized for ease of use and reliability.

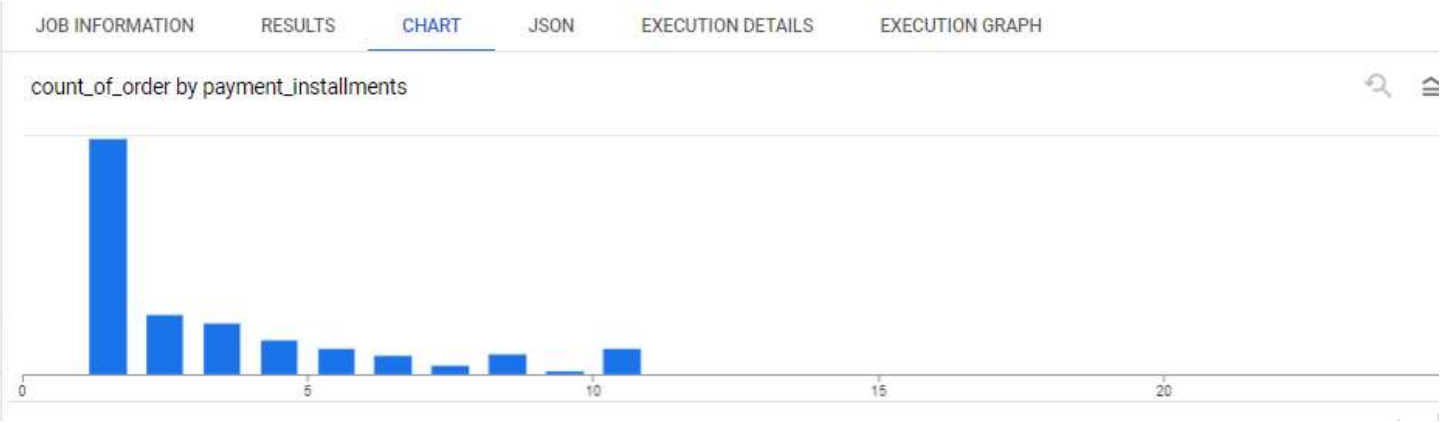
b. Find the no. of orders placed on the basis of the payment installments that have been paid :

QUERY:

```
select
payment_installments,
count(distinct order_id) as count_of_order
from `TARGET.Payments`
where payment_installments>=1
group by 1
order by 1
```

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	payment_installment	count_of_order
1	1	49060
2	2	12389
3	3	10443
4	4	7088
5	5	5234
6	6	3916
7	7	1623
8	8	4253
9	9	644
10	10	5315
11	11	23



INSIGHTS:

- **Good Payment Compliance:** The large number of orders with 1 or 2 installments remaining suggests that most customers were on track or nearly complete with their payments. This indicates good payment compliance and that the installment plans were manageable for most customers.
- **Low Default Risk:** Since the majority of orders are nearing completion (with low remaining installments), the risk of default or loss due to unpaid installments is relatively low. This suggests that the company was likely effective in structuring installment plans that matched customer capabilities.
- **Successful Installment Strategy:** The distribution of remaining installments suggests that the company's installment strategy was generally successful, with most customers keeping up with their payment schedules. The data shows that installment plans were largely accessible and appealing to customers.
- **Customer Affordability:** The trend indicates that customers had the financial ability to keep up with payments, as most are close to paying off their balances. This could be a sign that the company targeted the right customer segment with appropriate financing terms.
- **Potential Cash Flow Management Insight:** With most orders having low remaining installments, the company likely experienced a steady cash flow from installment payments. This regular income stream might have been important for managing operational expenses up to the point of closure.
- **Operational Closure Planning:** Since the data is from the end of the company's operations, the company might have been in a phase of winding down its receivables. The high completion rate of installments could suggest that the company was nearing a point where most of its accounts receivable were being settled.
- **Opportunity to Clear Dues:** With many customers having only 1 remaining installment, there may have been an opportunity for the company to clear dues quickly if required, through targeted communication or small incentives for final payments.

Overall, this analysis indicates that the company maintained a strong control over its installment payment plans, with the vast majority of customers managing to stay on schedule or close to completion by the end of operations.