Target SQL Business Case

-Saurabh Deshetti

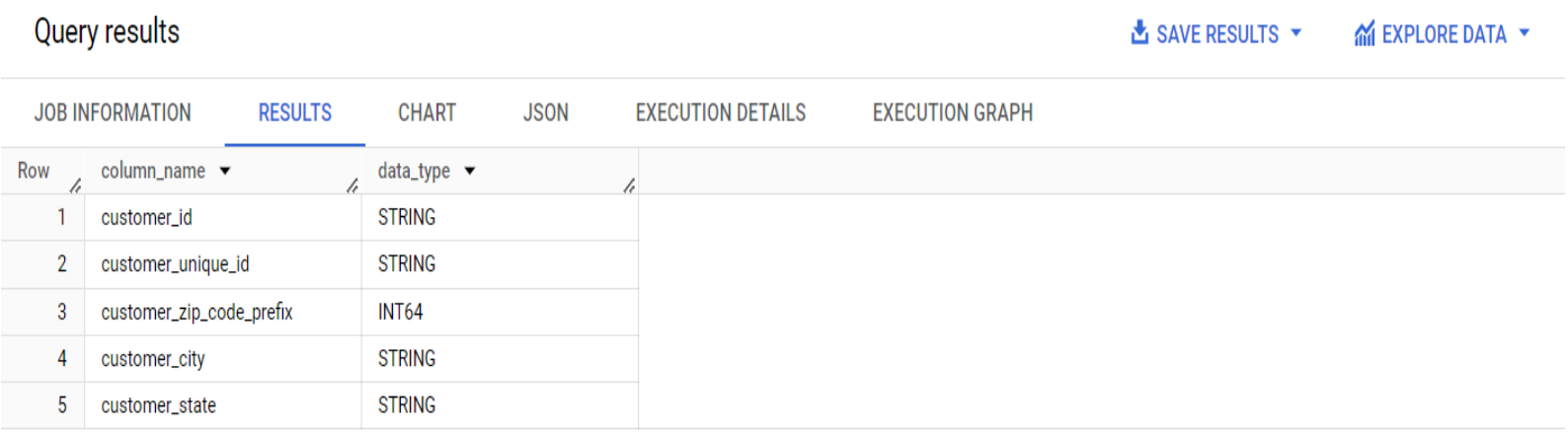
1. Import the dataset and do usual exploratory analysis steps like checking the structure & characteristics of the dataset:

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

- Query:

```
select column_name, data_type
from `Target_SQL.INFORMATION_SCHEMA.COLUMNS`
where table_name = 'customers'
```

- Query result:



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: In customers table,

1. The datatype of columns ‘customer_id’ is ‘STRING’.
2. The datatype of column ‘customer_unique_id’ is ‘STRING’.
3. The datatype of column ‘customer_zip_code_prefix is ‘INT64’.
4. The datatype of column ‘customer_city’ is ‘STRING’.
5. The datatype of column ‘customer_state’ is ‘STRING’.

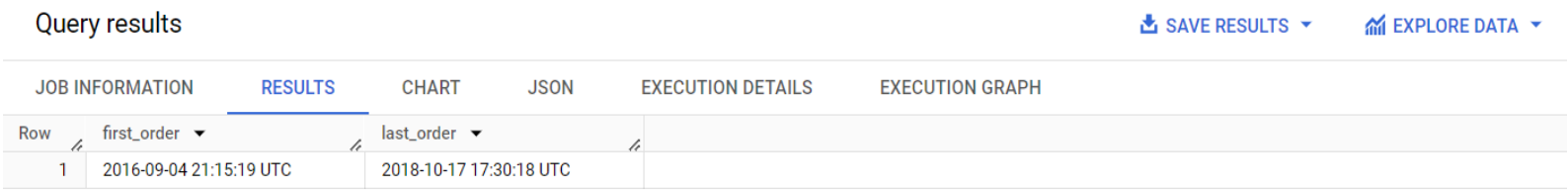
- Action items: NA

Q.2: Get the 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_SQL.orders`
```

- Query result:



- Insights: In the given dataset,

1. The first order was placed on 4th September 2016 at 21:15:19 UTC.

2. The last order was placed on 17th October 2018 at 17:30:18 UTC.

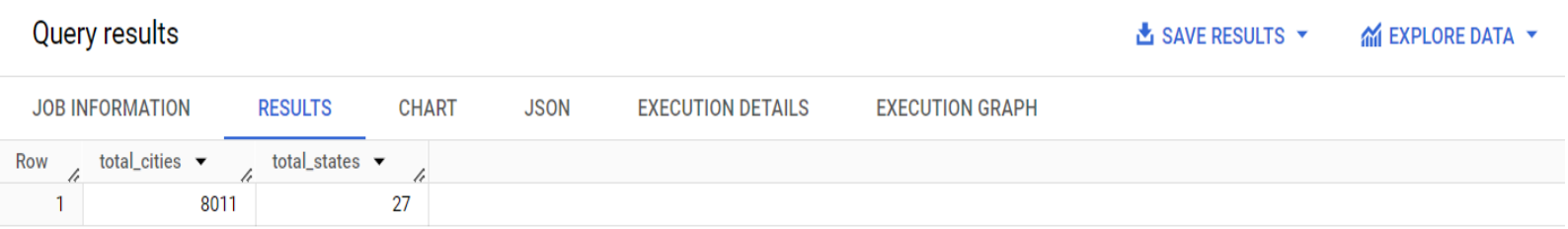
- Action items: NA

Q.3: Count the Cities & States of customers who ordered during the given period.

- Query:

```
select count(distinct geolocation_city) as total_cities,
count(distinct geolocation_state) as total_states
from `Target_SQL.geolocation`
```

- Query result:



- Insights: In the given dataset,

1. The number of unique cities are ‘8011’ & states are ‘27’ from where the orders were placed by the customers during the given time period.

- Action items: NA

2. In-depth Exploration:

Q.1: Is there a growing trend in the no. of orders placed over the past years?

- Query:

```
select order_year_month, count(order_purchase_timestamp) as
total_orders
from
(select order_purchase_timestamp,
FORMAT_DATE('%Y-%m',order_purchase_timestamp) AS
order_year_month
from `Target_SQL.orders`) tbl
group by order_year_month
order by order_year_month
```

- Query result:

Query results

SAVE RESULTS EXPLORE DATA

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

Row	order_year_month	total_orders
1	2016-09	4
2	2016-10	324
3	2016-12	1
4	2017-01	800
5	2017-02	1780
6	2017-03	2682
7	2017-04	2404
8	2017-05	3700
9	2017-06	3245
10	2017-07	4026

- Insights: Based on given dataset,

1. From overall observation it can be stated that the no. of orders placed has increased in each month, over the past years with some fluctuations in between.

(Exclude the initial and final two months from the observation for improved clarity.)

- Action items: According to my understanding,

1. The most important factors that affect directly to the sales of any e-commerce are:

    - Website navigation
    - Product range
    - Customer service
    - Number of sellers
    - Ease of ordering
    - Return and exchange policies
    - Trust over time
    - Delivery time
    - Multiple payment options

2. By monitoring and analysing these factors, e-commerce companies can identify opportunities for growth and implement strategies to optimize sales performance.

Q.2: Can we see some kind of monthly seasonality in terms of the no. of orders being placed?

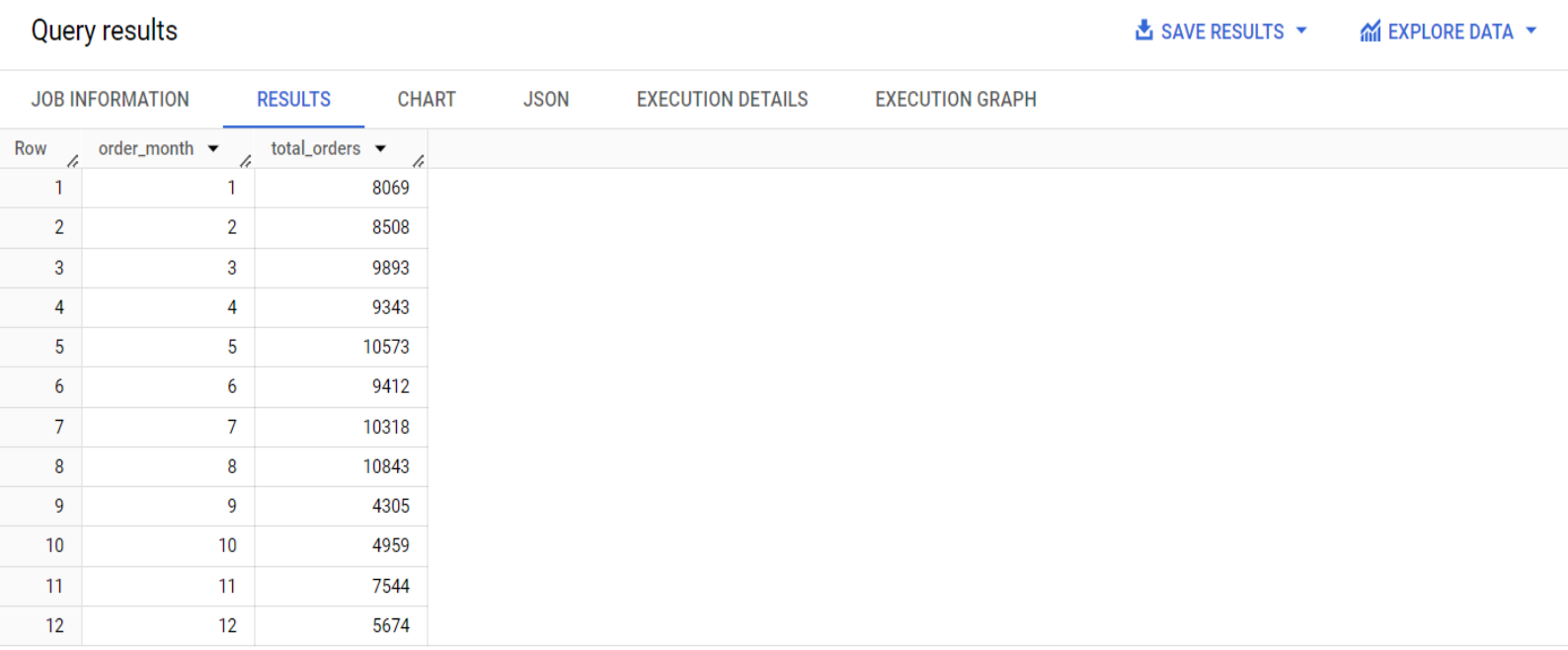
- Query:

```
select order_month,

count(order_purchase_timestamp) as total_orders

from
(select order_purchase_timestamp,
extract(MONTH from order_purchase_timestamp) as
order_month
from `Target_SQL.orders`) tbl
group by order_month
order by order_month
```

- Query result:



Row	order_month	total_orders
1	1	8069
2	2	8508
3	3	9893
4	4	9343
5	5	10573
6	6	9412
7	7	10318
8	8	10843
9	9	4305
10	10	4959
11	11	7544
12	12	5674

- Insights: Based on given dataset,

1. We can observe some monthly seasonality in no. of orders being placed.

2. The count of orders generally increases from March to August with some fluctuations in between.

3. August recorded the highest number of orders, followed by September which experienced the lowest order count immediately after the peak in order volume.

4. September to December consistently show the lowest order counts among all months.

- Action items: According to my understanding,

1. As we can observe the dip in order count from September to December, these are the months that should be targeted for any promotional offers or discount coupons.

2. Discount coupons and promotional offers are the best ways to attract customers during the off season for shopping over the e-commerce.

3. E-commerce companies can plan for a "Stock Clearance Sale" or an "End of Year Sale" to counter the declining sales trend at the end of the year.

Such initiatives can effectively capture the interest of customers who may have not bought anything during this period.

Q.3: During what time of the day, do the Brazilian customers mostly place their orders? (Dawn, Morning, Afternoon or Night)

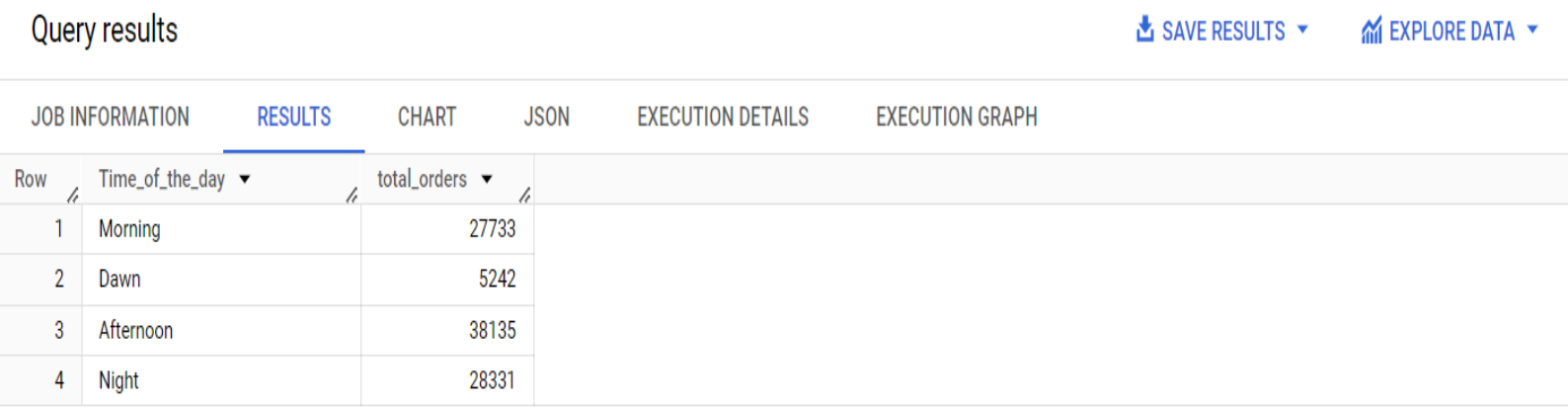
- 0-6 hrs: Dawn
- 7-12 hrs: Mornings
- 13-18 hrs: Afternoon,
- 19-23 hrs: Night

- Query:

```
select Time_of_the_day,

count(order_purchase_timestamp) as total_orders
from
(select order_purchase_timestamp,
case
when order_time between '00:00:00' and '06:59:59' then 'Dawn'
when order_time between '07:00:00' and '12:59:59' then 'Morning'
when order_time between '13:00:00' and '18:59:59' then 'Afternoon'
else 'Night'
end as Time_of_the_day
from
(select order_purchase_timestamp,
extract(time from order_purchase_timestamp) as order_time,
from `Target_SQL.orders`) tbl)tbl2
group by Time_of_the_day
```

- Query result:



Row	Time_of_the_day	total_orders
1	Morning	27733
2	Dawn	5242
3	Afternoon	38135
4	Night	28331

- Insights: Based on given dataset,

1. Customers tend to place the majority of their orders during the afternoon and evening hours. i.e. between 01 p.m. to 12 a.m.

2. Customers typically place fewer orders during the early morning hours, defined as between 12 a.m. to 7 a.m.

- Action items: According to my understanding,

1. By identifying off-peak purchasing period and peak buying times, e-commerce companies can work more efficiently on their resource allocation like,

    - Marketing campaigns
    - Promotional offers
    - Limited-Time Offers
    - Deploying customer service representatives
    - Managing inventory
    - Website Performance

2. This would definitely help E-commerce companies to optimise and use their resources in best possible ways, thereby maximizing their reach and potential sales opportunities.

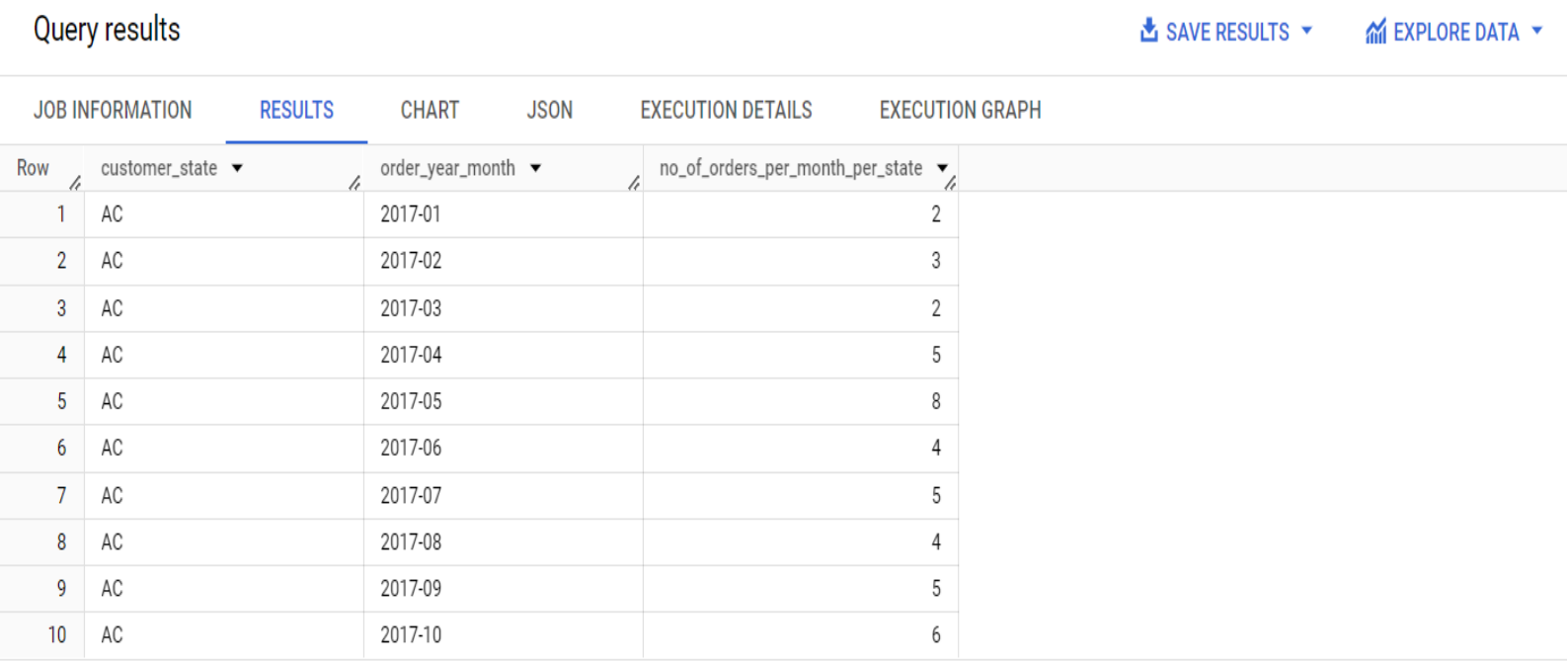
3. Evolution of E-commerce orders in the Brazil region:

Q.1: Get the month-on-month no. of orders placed in each state.

- Query:

```
select customer_state,
FORMAT_DATE('%Y-%m',order_purchase_timestamp) AS
order_year_month,
count(order_id) as no_of_orders_per_month_per_state
from `Target_SQL.customers` as c
join `Target_SQL.orders` as o
on c.customer_id = o.customer_id
group by customer_state, order_year_month
order by customer_state, order_year_month
```

- Query result:



Row	customer_state	order_year_month	no_of_orders_per_month_per_state
1	AC	2017-01	2
2	AC	2017-02	3
3	AC	2017-03	2
4	AC	2017-04	5
5	AC	2017-05	8
6	AC	2017-06	4
7	AC	2017-07	5
8	AC	2017-08	4
9	AC	2017-09	5
10	AC	2017-10	6

- Insights: Based on given dataset,

1.The sate ‘SP’ has the highest number of orders in all months.

2. The state ‘AC’ has the lowest number of orders in all months.

3. The month-on-month number of order count has gradually increased for most of the states with some fluctuations in between.

- Action items: According to my understanding,

1. There are some states like sate ‘SP’, state ‘MG’ and state ‘RJ’ have the maximum number of orders in all months.

These are the states that need more inventory and the most efficient customer service.

2. There are some states like state ‘AC’ and stat ‘AL’ have the minimum number of orders in all months.

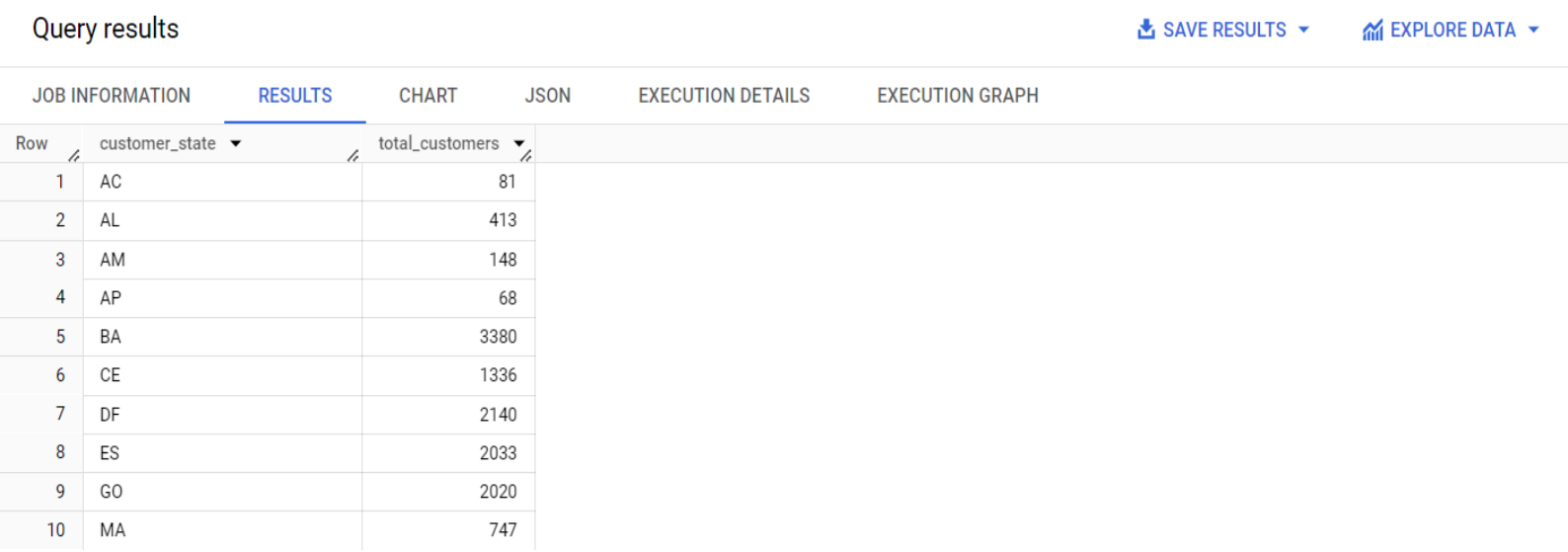
These are the states that need less inventory and can be targeted for marketing and promotional offers to attract more number customers and increase the sales.

Q.2: How are the customers distributed across all the states.

- Query:

```
select customer_state,
count(customer_unique_id) as total_customers
from `Target_SQL.customers`
group by customer_state
order by customer_state
```

- Query result:



Row	customer_state	total_customers
1	AC	81
2	AL	413
3	AM	148
4	AP	68
5	BA	3380
6	CE	1336
7	DF	2140
8	ES	2033
9	GO	2020
10	MA	747

- Insights: Based on given dataset,

1. The state ‘SP’ has the highest number of customers ‘41746’ followed by state ‘RJ’ with the number of customers ‘12852’ and state ‘MG’ with the number of customers ‘11635’.

2. The state ‘RR’ has the lowest number of customers ‘46’ followed by state ‘AP’ with the number of customers ‘68’ and state ‘AC’ with the number of customers ‘81’.

- Action items: According to my understanding,

1. By understanding number of customers in each state the e-commerce companies can focus on optimization of their resources like logistics, customer support and marketing strategies so that the resources can get used at their best ways possible and deliver personalized experience for customers in different states.

2. The states like ‘SP’, ‘RJ’ and ‘MG’ with highest number of customers definitely need more resources compared to all others states.

3. The states like ‘RR’, ‘AP’ and ‘AC’ with lowest number of customers definitely need less resources compared to all others states.

4. This would definitely help E-commerce companies to optimise and use their resources in best possible ways, thereby maximizing their reach and potential sales opportunities.

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

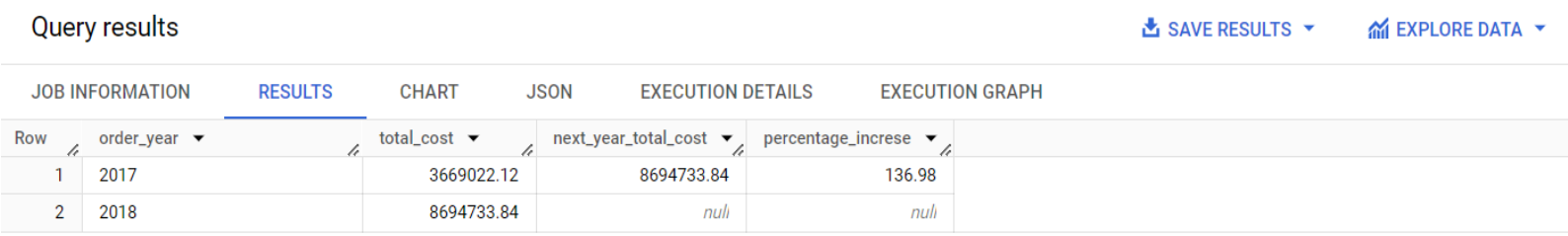
Q.1: Get the % increase in the cost of orders from year 2017 to 2018 (include months between Jan to Aug only).

You can use the "payment_value" column in the payments table to get the cost of orders.

- Query:

```
select *,
round(((next_year_total_cost - total_cost)/ total_cost *
100),2) as percentage_increse
from
(select *,
round(lead (total_cost) over(order by order_year),2) as
next_year_total_cost,
from
(select order_year, round(sum (payment_value),2) as
total_cost,
from
(select o.order_purchase_timestamp,
FORMAT_DATE('%Y',order_purchase_timestamp) AS order_year,
FORMAT_DATE('%m',order_purchase_timestamp) AS order_month,
p.payment_value
from `Target_SQL.payments` as p
join `Target_SQL.orders` as o
on p.order_id = o.order_id) tbl
where order_year in ('2017', '2018')
and order_month between '01' and '08'
group by order_year) tbl2
order by order_year) tbl3
```

- Query result:


Query results

SAVE RESULTS EXPLORE DATA

JOB INFORMATION RESULTS CHART JSON EXECUTION DETAILS EXECUTION GRAPH

Row	order_year	total_cost	next_year_total_cost	percentage_increse
1	2017	3669022.12	8694733.84	136.98
2	2018	8694733.84	null	null

- Insights: Based on given dataset,

1. The overall percentage increase in the cost of orders from year 2017 to 2018 (including only the months from January to August) is “136.98%”.

- Action items: According to my understanding,

1. Tracking the percentage increase in the cost of orders between any two years enables e-commerce businesses to monitor their growth and strategize for future campaigns and discount sales effectively.

2. E-commerce companies can target on factors that are responsible for increase or decrease in orders like,

- User experience
- Product quality and variety
- Pricing
- Supply chain and logistics

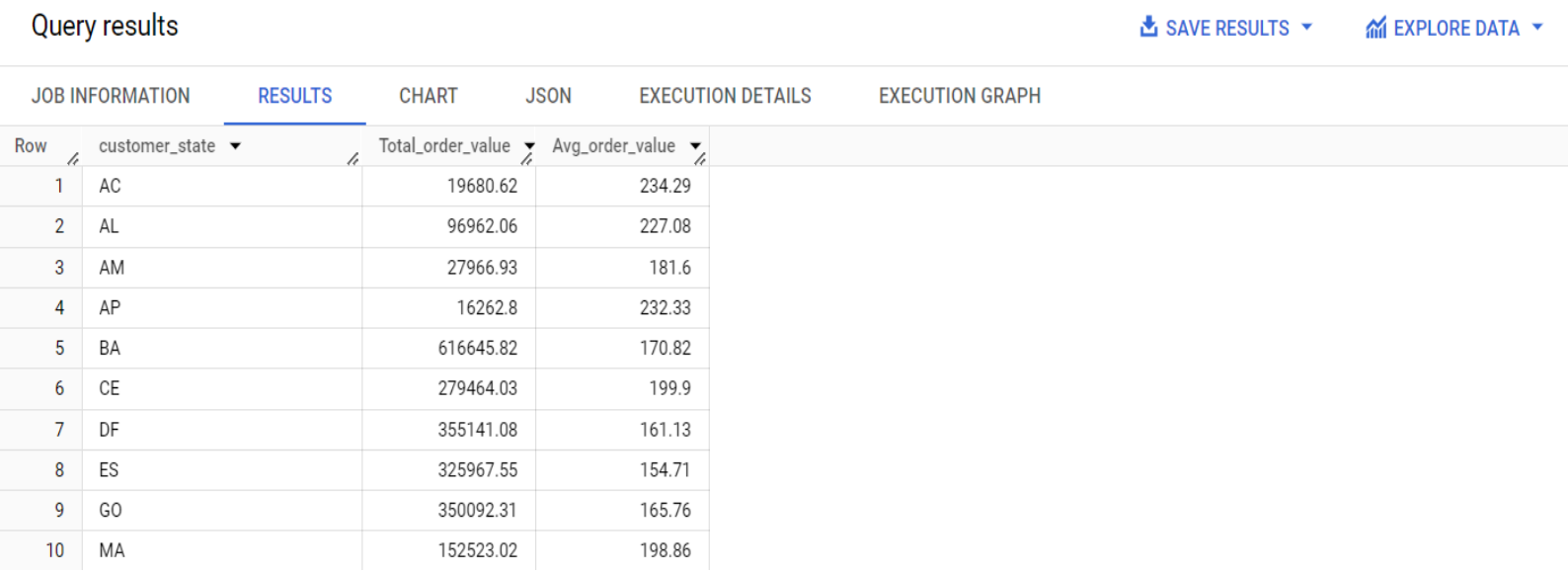
3.This would definitely help E-commerce companies to optimise and use their resources in best possible ways.

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

- Query:

```
select customer_state,
round(sum(payment_value),2) as Total_order_value,
round(avg(payment_value),2) as Avg_order_value
from `Target_SQL.customers` as c
join `Target_SQL.orders` as o
on c.customer_id = o.customer_id
join `Target_SQL.payments` as p
on o.order_id = p.order_id
group by customer_state
order by customer_state
```

- Query result:



Row	customer_state	Total_order_value	Avg_order_value
1	AC	19680.62	234.29
2	AL	96962.06	227.08
3	AM	27966.93	181.6
4	AP	16262.8	232.33
5	BA	616645.82	170.82
6	CE	279464.03	199.9
7	DF	355141.08	161.13
8	ES	325967.55	154.71
9	GO	350092.31	165.76
10	MA	152523.02	198.86

- Insights: Based on given dataset,

1. The state ‘SP’ has highest total order price of ‘5998226.96’ followed by state ‘RJ’ with total order price of ‘2144379.69’ and state ‘MG’ with total order price of ‘1872257.26’.

2. The state ‘RR’ has lowest total order price of ‘10064.62’ followed by state ‘AP’ with total order price of ‘16262.8’ and state ‘AC’ with total order price of ‘19680.62’.

3. The state ‘PB’ has highest average order value of ‘248.33’ followed by state ‘AC’ with total order price of ‘234.29’ and state ‘RO’ with total order price of ‘233.2’.

4. The state ‘SP’ has lowest average order value of ‘137.5’ followed by state ‘PR’ with total order price of ‘154.15’ and state ‘MG’ with total order price of ‘154.71’.

- Action items: According to my understanding,

1. Through the examination of state-wise patterns, e-commerce companies can gain valuable insights into regional variations.

2.There are some state which have highest total order price but the lowest average order value like state ‘SP’ and state ‘MG’.

In such cases e-commerce companies can target those states for discount offers and gift vouchers to increase average order value for those states as those sates already having most numbers of customers.

3. There are some state which have highest average order vale but the lowest total order price like state ‘AC’.

In such cases e-commerce companies can target those states for marketing campaigns to gain customer base as the customers over those states have high spending capacities.

4. This analytical approach helps e-commerce companies to find hidden opportunities for growth and expansion.

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

- Query:

```
select customer_state,

round(sum(freight_value),2) as Total_order_freight,
round(avg(freight_value),2) as Avg_order_freight
from `Target_SQL.customers` as c
join `Target_SQL.orders` as o
on c.customer_id = o.customer_id
join `Target_SQL.order_items` as oi
on o.order_id = oi.order_id
group by customer_state
order by customer_state
```

- Query result:

Query results

SAVE RESULTS EXPLORE DATA

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

Row	customer_state	Total_order_freight	Avg_order_freight
1	AC	3686.75	40.07
2	AL	15914.59	35.84
3	AM	5478.89	33.21
4	AP	2788.5	34.01
5	BA	100156.68	26.36
6	CE	48351.59	32.71
7	DF	50625.5	21.04
8	ES	49764.6	22.06
9	GO	53114.98	22.77
10	MA	31523.77	38.26

- Insights: Based on given dataset,

1. The state ‘SP’ has highest value of total order freight of ‘718723.07’ followed by state ‘RJ’ with total order freight of ‘305589.31’ and state ‘MG’ with total order freight of ‘270853.46’.

2. The state ‘RR’ has lowest value of total order freight of ‘2235.19’ followed by state ‘AP’ with total order freight of ‘2788.5’ and state ‘AC’ with total order freight of ‘3686.75’.

3. The state ‘RR’ has highest average order freight of ‘42.98’ followed by state ‘PB’ with total order freight of ‘42.72’ and state ‘RO’ and state ‘AC’ with total order freight of ‘41.07’.

4. The state ‘SP’ has lowest average order freight of ‘15.15’ followed by state ‘PR’ with total order freight of ‘20.53’ and state ‘MG’ with total order freight of ‘20.63’.

- Action items: According to my understanding,

1. By the use of data-driven insights, e-commerce companies can optimize their pricing strategies and their logistics to provide personalized experience.

2.There are some state which have lowest value of total order freight but the highest average order freight like state ‘RR’ and state ‘AC’.

In such cases e-commerce companies can target those states for reducing average order freight so that more customers can place orders without thinking about freight value so that they can experience good value for money proposition.

3. Ultimately, by using state-wise analyses, e-commerce companies can position themselves in better way.

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

Q.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.

Do this in a single query.

You can calculate the delivery time and the difference between the estimated & actual delivery date using the given formula:

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

- Query:

```
select order_id,

date_diff(order_delivered_customer_date,
order_purchase_timestamp, day) as delivery_time,
date_diff(order_estimated_delivery_date,
order_delivered_customer_date, day) as
diff_estimated_delivery
from `Target_SQL.orders`
where order_delivered_customer_date is not null
order by order_id
```

- Query result:

Query results

SAVE RESULTS EXPLORE DATA

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

Row	order_id	delivery_time	diff_estimated_delivery
1	00010242fe8c5a6d1ba2dd792...	7	8
2	00018f77f2f0320c557190d7a1...	16	2
3	000229ec398224ef6ca0657da...	7	13
4	00024acbcdf0a6daa1e931b03...	6	5
5	00042b26cf59d7ce69dfabb4e...	25	15
6	00048cc3ae777c65dbb7d2a06...	6	14
7	00054e8431b9d7675808bcb8...	8	16
8	000576fe39319847cbb9d288c...	5	15
9	0005a1a1728c9d785b8e2b08...	9	0
10	0005f50442cb953dcd1d21e1f...	2	18

- Insights: Based on given dataset,

1. Minimum days taken to deliver any order is ‘0’ days.

It means the order was delivered within 24hrs from the time of order being placed.

2. Maximum days taken to deliver any order is ‘209’ days.

It means the order took 209 days to deliver from the date of order being placed.

3. The order which is delivered fastest was ‘146’ days before estimated delivery date.

4. The order which took too long to deliver was ‘188’ days after estimated delivery date.

- Action items: According to my understanding,

1. Orders which delivered before estimated date of delivery won't make any issue but the orders which delivered after estimated date of delivery may affect the trust of customers which is one of the most important factors for any e-commerce company.

2. E-commerce companies to focus on the orders which are being delivered after estimated date of delivery and try to deliver them within estimated date of delivery or try providing exact estimated date of delivery at the time of order being placed.

3. The orders delivered within estimated date of delivery will help in enhancing customer experience and will lead to the growth in sales of any e-commerce company.

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

- Query:

```
Select customer_state, avg_freight_value

From
(select customer_state, avg_freight_value,
row_number() over(order by avg_freight_value) top_5_lowest,
row_number() over(order by avg_freight_value desc)
top_5_highest
from
(select customer_state,
round(avg(freight_value),2) as avg_freight_value,
from `Target_SQL.customers` as c
join `Target_SQL.orders` as o
on c.customer_id = o.customer_id
join `Target_SQL.order_items` as oi
on o.order_id = oi.order_id
group by customer_state
order by avg_freight_value) tbl)tbl2
Where top_5_highest <=5 or top_5_lowest <=5
order by avg_freight_value
```

- Query result:

Row	customer_state	avg_freight_value
1	SP	15.15
2	PR	20.53
3	MG	20.63
4	RJ	20.96
5	DF	21.04
6	PI	39.15
7	AC	40.07
8	RO	41.07
9	PB	42.72
10	RR	42.98

- Insights: Based on given dataset,

1. The state ‘SP’ has lowest average freight value of ‘15.15’ followed by state ‘PR’ with average freight value of ‘20.53’ and state ‘MG’ with average freight value of ‘20.63’.

2. The state ‘RR’ has highest average freight value of ‘42.98’ followed by state ‘PB’ with average freight value of ‘42.72’ and state ‘RO’ with average freight value of ‘41.07’.

- Action items: According to my understanding,

1. The states where average freight value is highest like state ‘RR’, state ‘PB’, state ‘RO’, state ‘AC’ and state ‘PI’ should be worked on comparing those with states having lowest average freight value.

2. The states where average freight value is highest can impact on sales due to the freight value. It can make customers to think twice before placing any order due to high impact over value for money proposition.

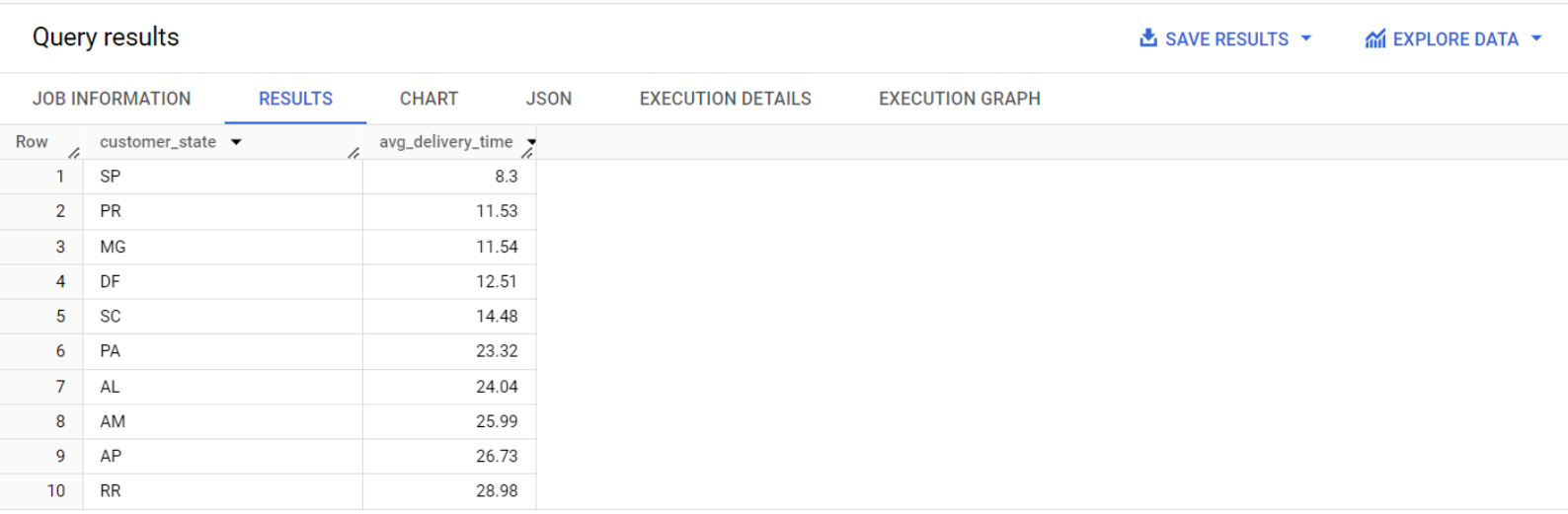
Q.3: Find out the top 5 states with the highest & lowest average delivery time.

- Query:

```
Select customer_state, avg_delivery_time

From
(select customer_state, avg_delivery_time,
row_number() over(order by avg_delivery_time) top_5_lowest,
row_number() over(order by avg_delivery_time desc)
top_5_highest
from
(select customer_state,
round(avg(date_diff(order_delivered_customer_date,
order_purchase_timestamp, day)),2) as avg_delivery_time,
from `Target_SQL.customers` as c
join `Target_SQL.orders` as o
on c.customer_id = o.customer_id
group by customer_state
order by avg_delivery_time) tbl) tbl2
Where top_5_highest <=5 or top_5_lowest <=5
order by avg_delivery_time
```

- Query result:



Row	customer_state	avg_delivery_time
1	SP	8.3
2	PR	11.53
3	MG	11.54
4	DF	12.51
5	SC	14.48
6	PA	23.32
7	AL	24.04
8	AM	25.99
9	AP	26.73
10	RR	28.98

- Insights: Based on given dataset,

1. The state ‘SP’ has lowest average delivery time of ‘8.3’ days followed by state ‘PR’ with average delivery time of ‘11.53’ days and state ‘MG’ with average delivery time of ‘11.54’ days.

2. The state ‘RR’ has highest average delivery time of ‘28.98’ days followed by state ‘AP’ with average delivery time of ’26.73’ days and state ‘AM’ with average delivery time of ’25.99’ days.

- Action items: According to my understanding,

1.The states with highest average delivery time like state ‘RR’, state ‘AP’, state ‘AM’, state ‘AL’, and state ‘PA’ should be worked on to improve delivery time.

2. Any e-commerce company can improve their tame to complete any delivery by just working on logistics and warehouse of those particular states where they need to improve their average delivery time.

Q.4: Find out the top 5 states where the order delivery is really fast as compared to the estimated date of delivery.

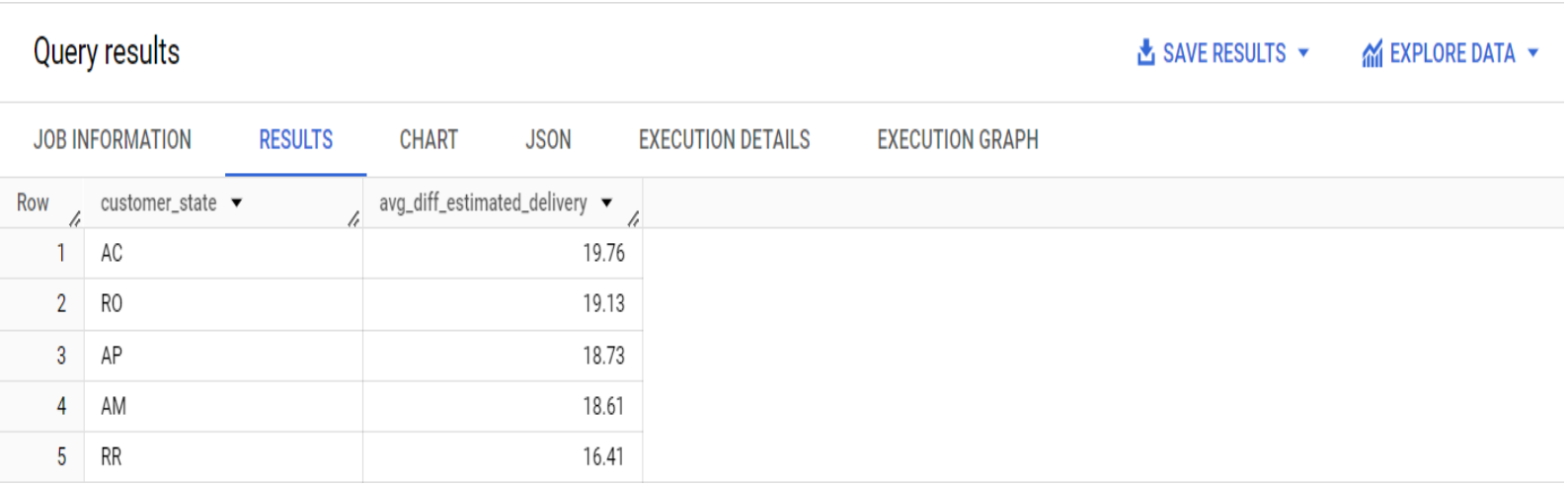
You can use the difference between the averages of actual & estimated delivery date to figure out how fast the delivery was for each state.

- Query:

```
select customer_state,

round(avg(diff_estimated_delivery),2) as
avg_diff_estimated_delivery
from
(select order_id, customer_state,
date_diff(order_estimated_delivery_date,
order_delivered_customer_date,day) as diff_estimated_delivery
from `Target_SQL.customers` as c
join `Target_SQL.orders` as o
on c.customer_id = o.customer_id
where order_delivered_customer_date is not null)tbl
group by customer_state
order by avg_diff_estimated_delivery desc
limit 5
```

- Query result:



Row	customer_state	avg_diff_estimated_delivery
1	AC	19.76
2	RO	19.13
3	AP	18.73
4	AM	18.61
5	RR	16.41

- Insights: Based on given dataset,

1. In the state ‘AC’ the average order delivery is really fast compared to estimated date of delivery with the average of ’19.67’ days before the estimated date of delivery followed by state ‘RO’ with average of ’19.13’ days before the estimated date of delivery and state ‘AP’ with average of ‘18.73’ days before the estimated date of delivery.

- Action items: According to my understanding,

1. The states where the average order delivery is really fast compared to estimated date of delivery can be targeted for marketing campaigns using delivery as selling point for all the products in those particular states.

2. By monitoring and analysing these factors, e-commerce companies can identify opportunities for growth, address challenges, and implement strategies to optimize sales performance.

6. Analysis based on the payments:

Q.1: Find the month-on-month no. of orders placed using different payment types.

- Query:

```
select FORMAT_DATE('%Y-%m', order_purchase_timestamp)
as year_month,

payment_type,
count(o.order_id) as total_orders
from `Target_SQL.orders` as o
join `Target_SQL.payments` as p
on o.order_id = p.order_id
group by year_month, payment_type
order by year_month, payment_type
```

- Query result:

Query results

SAVE RESULTS | EXPLORE DATA

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

Row	year_month	payment_type	total_orders
1	2016-09	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	2017-01	UPI	197
8	2017-01	credit_card	583
9	2017-01	debit_card	9
10	2017-01	voucher	61

- Insights: Based on given dataset,

1. Credit card are used by the most people as a payment method, followed by UPI.

2. Debit card are used by the least people as a payment method.

3. Most of the people using credit card as payment method due to the benefits of using credit cards for shopping like cashback or reward vouchers.

- Action items: According to my understanding,

1. As we can observe most people use Credit cards for shopping due to benefits of using credit cards.

E-commerce companies can try providing more offers for the customers using credit cards for shopping or e-commerce companies can launch features like ‘Pay later’ to attract more customers.

Q.2: Find the no. of orders placed on the basis of the payment instalments that have been paid.

- Query:

```
select p.payment_installments,

count(o.order_id) as total_orders
from `Target_SQL.orders` as o
join `Target_SQL.payments` as p
on o.order_id = p.order_id
where payment_installments != 0
and payment_value != 0
group by p.payment_installments
order by p.payment_installments
```

- Query result:

Query results

SAVE RESULTS | EXPLORE DATA

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

Row	payment_installment	total_orders
1	1	52537
2	2	12413
3	3	10461
4	4	7098
5	5	5239
6	6	3920
7	7	1626
8	8	4268
9	9	644
10	10	5328

- Insights: Based on given dataset,

1. The maximum number of orders have only one payment instalment this may due to high usage of credit cards for shopping.

2. The highest number of instalments is ‘24’ and is for total of ‘18’ orders.

- Action items: According to my understanding,

1. As we can observe most of the people are using single instalment for payment that may be due to high usage of credit cards so the e-commerce companies can offer vouchers and discounts around credit cards for engaging more customers and by providing flexibility in payment by offering more payment instalments to attract higher customer base.

2. In summary, the analysis gives us valuable insights regarding instalment preferences used by customers. It highlights the high usage of credit card transactions.