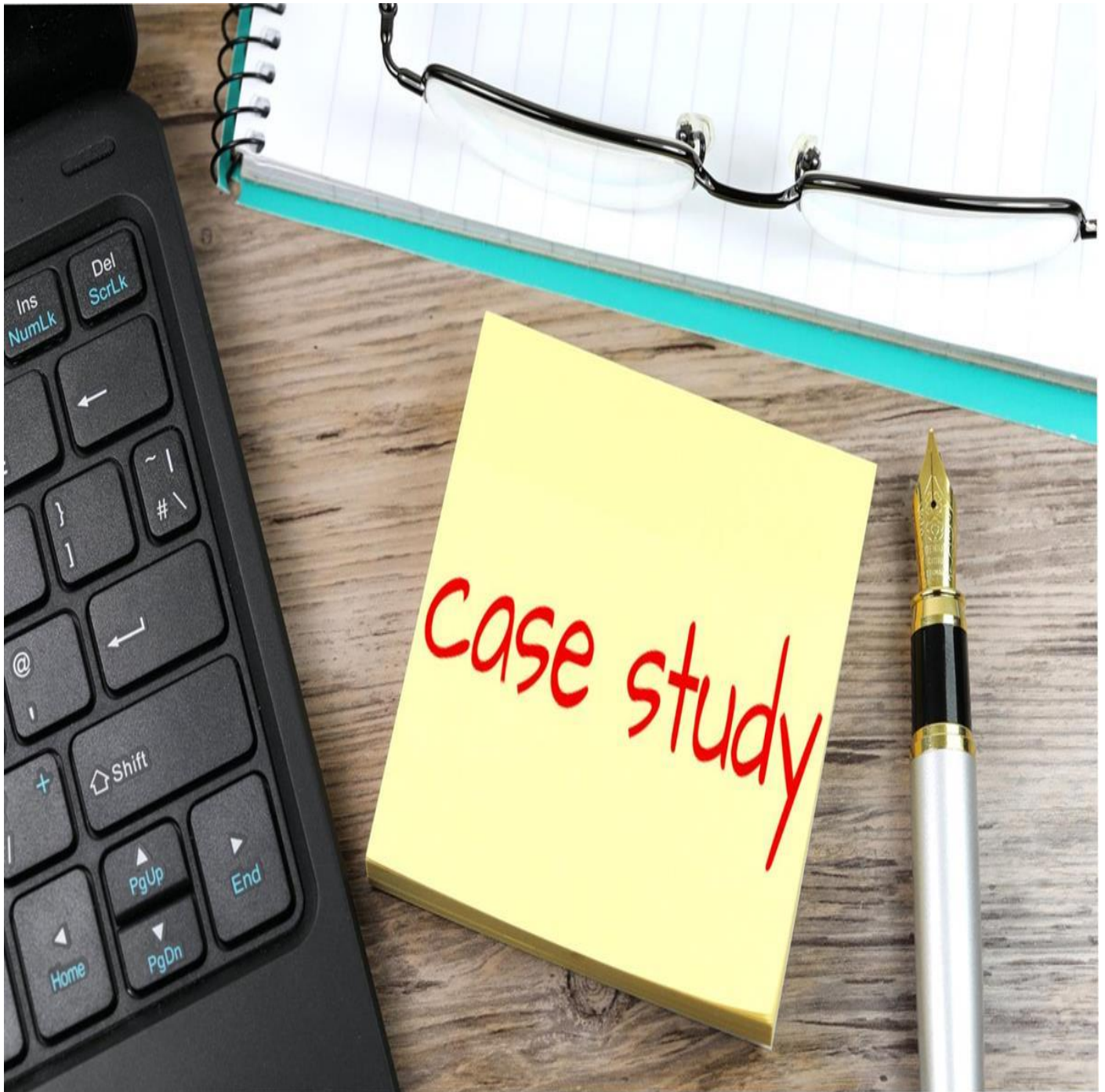


# **TARGET BUSINESS CASE**

**-BY SADHANA SUMAN**



## BUSINESS CASE: TARGET SQL

### DESCRIPTION:

Target Corporation is an American retail corporation. The company's headquarter is located in Minneapolis, Minnesota. It is the seventh largest retailer in the United States. Target's CEO since August 2014 is Brian Cornell. The goal is to make every family's Target Run convenient, relevant, affordable and packed with joyful experiences they won't find anywhere else.

Target has expanded its e-commerce operations in Brazil to grow online shopping trend. In order to ensure success in this competitive market, it is crucial for Target to understand the dynamics of e-commerce in Brazil and leverage data-driven insights to enhance its operations. In this case study, we will analyse the company's e-commerce dataset using structured query language (SQL) queries to provide recommendations for improving their operations in Brazil.

This case study 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.

After analysing this 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.

### PROBLEM STATEMENT:

Assuming you are a data analyst/ scientist at Target, you have been assigned the task of analysing the given dataset to extract valuable insights and provide actionable recommendations.

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

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

Before performing the analysis of the dataset, we started with verifying the column data types in the customers' table using the following SQL query.

QUERY:

```
SELECT
  COLUMN_NAME,
  DATA_TYPE
FROM
  farmers_market.INFORMATION_SCHEMA.COLUMNS
WHERE
  table_name='customer';
```

OUTPUT:

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	COLUMN_NAME ▼	DATA_TYPE ▼			
1	customer_id	INT64			
2	customer_first_name	STRING			
3	customer_last_name	STRING			
4	customer_zip	INT64			

INSIGHTS:

1) We performed the SQL query and found out that the customer\_first\_name column and customer\_last\_name have the data type 'string' while 'customer\_id' and 'customer\_zip' are in the form of integer.

2) This data is helpful in ensuring accurate analysis and interpretation of dataset.

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

To get the time range between the orders that were placed in the Brazil between the year 2016 and 2018 we performed the following query-

QUERY:

```
WITH
  cte AS (
  SELECT
```

```

*,
EXTRACT (date
FROM
    order_purchase_timestamp) AS order_date
FROM
    `target_sql.orders`)
SELECT
    DATE_DIFF(MAX(order_date),MIN(order_date),year) AS range_in_years,
    DATE_DIFF(MAX(order_date),MIN(order_date),month) AS range_in_months,
    DATE_DIFF(MAX(order_date),MIN(order_date),day) AS range_in_days
FROM
    cte;

```

OUTPUT:

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	//	range_in_years ▼	//	range_in_months ▼	//
1		2		25	773

INSIGHTS:

- 1)The time range in terms of years was found to be 2 years.
- 2) The time range for months was found to be 25 months.
- 3) The time range based on number of days was found to be 773 days.

1.3)Count the number of Cities and States in our dataset.

Before analysing the dataset, we performed a query to find the number of cities and states of customers who placed an order in the given time period in the Brazil.

QUERY:

```

SELECT
    COUNT(DISTINCT customer_city) AS total_city,
    COUNT(DISTINCT customer_state) AS total_state
FROM
    `target_sql.customers`;

```

OUTPUT:

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	//	total_city ▼	//	total_state ▼	//
1		4119		27	

INSIGHTS:

After analysing the count of cities and states in Brazil, we found out that there are 4119 cities and 27 states from where the customers ordered from the year 2016 to 2018 in our dataset.

Q.2) In-depth Exploration.

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

To analyse the growing trend in the number of orders we extracted the year, month and distinct no. of orders over the given time period.

QUERY:

```
SELECT
  EXTRACT(year
FROM
  order_purchase_timestamp) AS Year,
  EXTRACT(month
FROM
  order_purchase_timestamp) AS Month,
  COUNT(DISTINCT order_id) AS Total_orders
FROM
  `target_sql.orders`
GROUP BY
  Year,
  Month
ORDER BY
  Year,
  Month;
```

OUTPUT:

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	Year	Month	Total_orders		
1	2016	9	4		
2	2016	10	324		
3	2016	12	1		
4	2017	1	800		
5	2017	2	1780		
6	2017	3	2682		
7	2017	4	2404		
8	2017	5	3700		
9	2017	6	3245		
10	2017	7	4026		



## INSIGHTS:

1)After analysing the number of orders over the years and months we can conclude that there are some fluctuations but the overall trend is upwards.

2)Therefore based on the above analysis of order count, it can be observed that there is a growing trend in e-commerce in Brazil.

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

To determine the monthly seasonality, we extracted months and number of orders being placed in Brazil over the period of 2016-2018.

## QUERY:

```
SELECT
  EXTRACT(month
FROM
  order_purchase_timestamp) AS Month,
  COUNT(order_id) AS number_of_orders
FROM
  `target_sql.orders`
GROUP BY
  month
ORDER BY
  month;
```

## OUTPUT:

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	Month	number_of_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			



#### INSIGHTS:

- 1) From the above analysis we can observe some seasonality in the e-commerce orders.
- 2) We can see that count of orders generally increases from March to August but there are some fluctuations too in between.
- 3) There is an increase in orders during February and March, coinciding with the Carnival season in Brazil.
- 4) We can see peak in the month of August in order count, potentially related to the Festival de Cachaça dedicated to the national liquor, cachaça.

2.3) 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

We executed the following query to find out at what time of the day customers order the most. We categorize the order purchase timestamps into four periods: dawn, morning, afternoon, and night.

#### QUERY:

```
SELECT
CASE
WHEN EXTRACT(hour FROM order_purchase_timestamp) BETWEEN 0 AND 6 THEN 'Dawn'
WHEN EXTRACT(hour FROM order_purchase_timestamp) BETWEEN 7 AND 12 THEN 'Morning'
WHEN EXTRACT(hour FROM order_purchase_timestamp) BETWEEN 13 AND 18 THEN 'Afternoon'
WHEN EXTRACT(hour FROM order_purchase_timestamp) BETWEEN 19 AND 23 THEN 'Night'
```



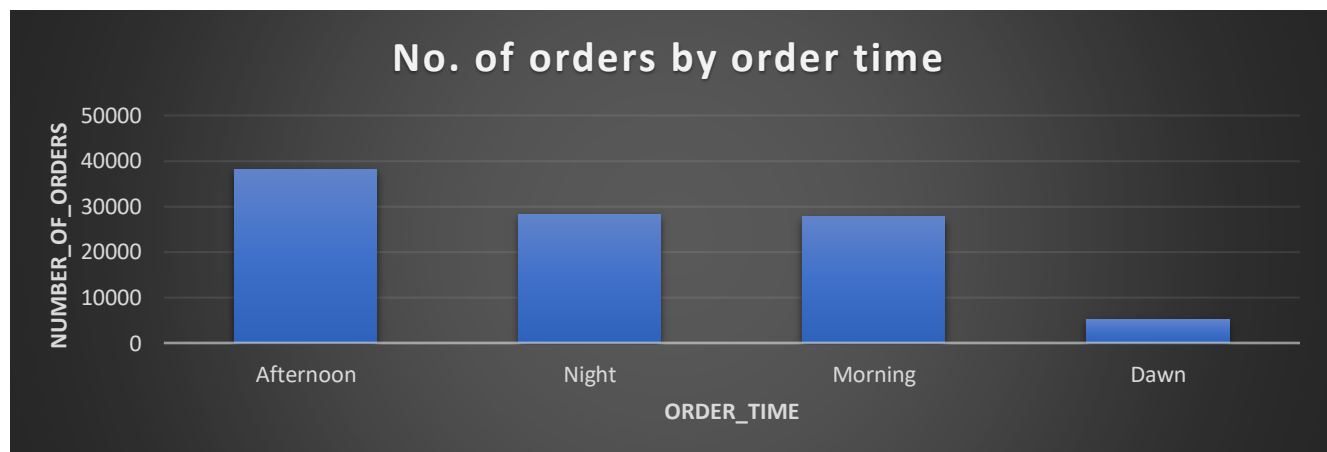
```

END
AS order_time,
COUNT(*) AS number_of_orders
FROM
`target_sql.orders`
GROUP BY
order_time
ORDER BY
number_of_orders DESC;

```

OUTPUT:

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	order_time	number_of_orders			
1	Afternoon	38135			
2	Night	28331			
3	Morning	27733			
4	Dawn	5242			



INSIGHTS:

- 1) From the above results we can clearly see that customers of Brazil orders the most during the afternoon and night. The total number of orders placed in the afternoon were found to be 38135.
- 2) We can say that customers shop during their free time.
- 3) After identifying peak buying times, companies can allocate resources, such as customer service representatives and inventory, more effectively to meet customer demands and provide a seamless shopping experience.

### Q.3) Evolution OF E-commerce orders in the Brazil region.

3.1) Get the month on month number of orders placed in each state.

To determine the evolution of e-commerce orders in the Brazil , we analysed the month-on-month order counts for each of the 27 states.

QUERY:

```
SELECT
  customer_state,
  EXTRACT(month
FROM
  order_purchase_timestamp) AS Month,
  COUNT(*) AS Total_orders,
FROM
  `target_sql.orders` o
INNER JOIN
  `target_sql.customers` c
ON
  o.customer_id=c.customer_id
GROUP BY
  customer_state,
  Month
ORDER BY
  1,
  2;
```

OUTPUT:

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS		EXECUTION GRAPH
Row	customer_state	Month	Total_orders			
1	AC	1	8			
2	AC	2	6			
3	AC	3	4			
4	AC	4	9			
5	AC	5	10			
6	AC	6	7			
7	AC	7	9			
8	AC	8	7			
9	AC	9	5			
10	AC	10	6			

INSIGHTS:

- 1) The above result shows the month-on-month order counts in each state of Brazil, providing insights into the customer purchase trends on a state-by-state basis.
- 2) We can clearly see that São Paulo (SP) consistently has the highest number of orders in any given month, followed by Rio de Janeiro (RJ) and Minas Gerais (MG).

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

We executed a SQL query to explore the e-commerce landscape in Brazil, for this we determined the number of customers in each of the 27 states.

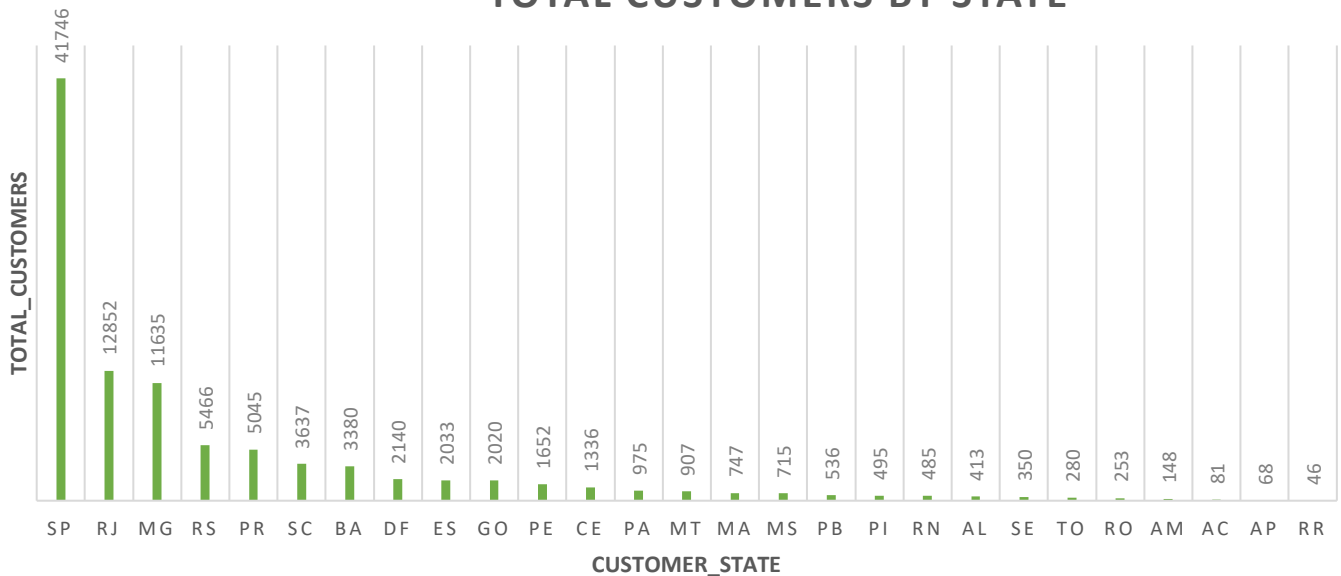
QUERY:

```
SELECT
  customer_state,
  COUNT(DISTINCT customer_id) AS total_customers,
FROM
  `target_sql.customers`
GROUP BY
  customer_state
ORDER BY
  total_customers DESC;
```

OUTPUT:

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	customer_state	total_customers			
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			

## TOTAL CUSTOMERS BY STATE



### INSIGHTS:

- 1)The above results reveal that the state of São Paulo (SP) has the highest number of customers that is 41746, which can be attributed to its status as the most populous state in Brazil. This finding also aligns with the previous analysis, indicating a positive correlation between the population of a state and its order count.
- 2)The understanding of evolution of e-commerce orders and the distribution of customers across Brazilian states is crucial in order to improve their business, customer experiences and market strategies.

Q4)Impact on Economy: Analyse the money movement by e-commerce by looking at order prices, freight and others.

4.1)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.

To determine the impact on the economy, we calculated the percentage increase in the cost of orders from 2017 to 2018, considering only the months from January to August. The following SQL query was executed:

### QUERY:

```
WITH
cte AS (
SELECT
    EXTRACT(year
FROM
    order_purchase_timestamp) AS Year,
    ROUND(SUM(payment_value),2) AS total_payment_value
FROM
    `target_sql.orders` o
INNER JOIN
    `target_sql.payments` p
ON
```

```

o.order_id=p.order_id
WHERE
EXTRACT(year
FROM
order_purchase_timestamp) IN(2017,
2018)
AND EXTRACT(month
FROM
order_purchase_timestamp) BETWEEN 1
AND 8
GROUP BY
year
ORDER BY
year)
SELECT
ROUND(((tb.lead_value-tb.total_payment_value)*100/tb.total_payment_value),2) AS growth_rate
FROM (
SELECT
year,
total_payment_value,
LEAD(total_payment_value) OVER (ORDER BY cte.year ASC) AS lead_value
FROM
cte
ORDER BY
year ASC) AS tb;

```

OUTPUT:

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	growth_rate ▼				
1	136.98				
2	null				

INSIGHTS:

1)From the payments table from our dataset we analysed the percentage increase in the cost of orders from year 2017 to 2018 and it was found to be Approximately 137 % from the months between January to August.

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

To determine the price and freight values on a state level, we calculated the average and sum of these values by a customer state. The following SQL query was executed

QUERY:

```
SELECT
    c.customer_state,
    ROUND(SUM(price),2) AS Total_price,
    ROUND(AVG(price),2) AS Avg_price
FROM
    `target_sql.orders` o
INNER JOIN
    `target_sql.order_items` oi
ON
    o.order_id=oi.order_id
INNER JOIN
    `target_sql.customers` c
ON
    o.customer_id=c.customer_id
GROUP BY
    customer_state
ORDER BY
    2 DESC,
    3 ;
```

OUTPUT:

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS		EXECUTION GRAPH
Row		customer_state	Total_price	Avg_price		
1		SP	5202955.05	109.65		
2		RJ	1824092.67	125.12		
3		MG	1585308.03	120.75		
4		RS	750304.02	120.34		
5		PR	683083.76	119.0		
6		SC	520553.34	124.65		
7		BA	511349.99	134.6		
8		DF	302603.94	125.77		
9		GO	294591.95	126.27		
10		ES	275037.31	121.91		



## INSIGHTS:

- 1) From the above results we can clearly see that the state SP has the highest total price value while the average price value for it is lowest.
- 2) For understanding the impact of economy the analysis of total and average price value is utmost important for growing the business, making market and pricing strategies.

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

We have calculated the total and average value of order price for each state through the following SQL query:

## QUERY:

```
SELECT
  c.customer_state,
  ROUND(SUM(freight_value),2) AS Total_freight,
  ROUND(AVG(freight_value),2) AS Avg_freight
FROM
  `target_sql.orders` o
INNER JOIN
  `target_sql.order_items` oi
ON
  o.order_id=oi.order_id
INNER JOIN
  `target_sql.customers` c
ON
  o.customer_id=c.customer_id
```

GROUP BY

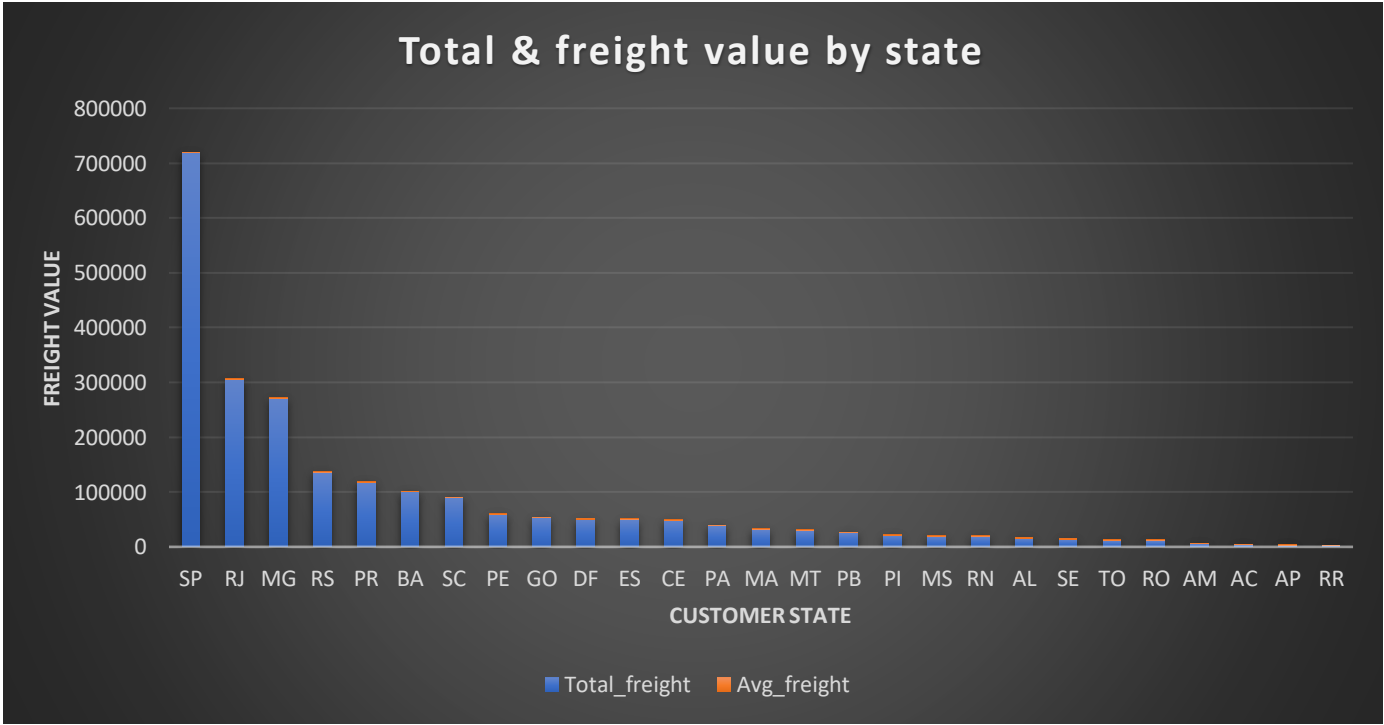
customer\_state

ORDER BY

2 desc,  
3 ;

OUTPUT:

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	customer_state	Total_freight	Avg_freight		
1	SP	718723.07	15.15		
2	RJ	305589.31	20.96		
3	MG	270853.46	20.63		
4	RS	135522.74	21.74		
5	PR	117851.68	20.53		
6	BA	100156.68	26.36		
7	SC	89660.26	21.47		
8	PE	59449.66	32.92		
9	GO	53114.98	22.77		
10	DF	50625.5	21.04		





## INSIGHTS:

- 1) From the results we can observe that the state SP has the highest total freight value that is a price at which a certain cargo is delivered from one point to another while it has the lowest average freight value.
- 2) On the other hand the state RR has the highest average freight value.
- 3) The analysis of freight value is important for economic growth, delivery of products and their distribution. Therefore, by analysing it we can take care of the business well.

## Q5) Analysis based on sales, freight and delivery time.

5.1) 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:

$$\text{time\_to\_deliver} = \text{order\_delivered\_customer\_date} - \text{order\_purchase\_timestamp}$$
$$\text{diff\_estimated\_delivery} = \text{order\_estimated\_delivery\_date} - \text{order\_delivered\_customer\_date}.$$

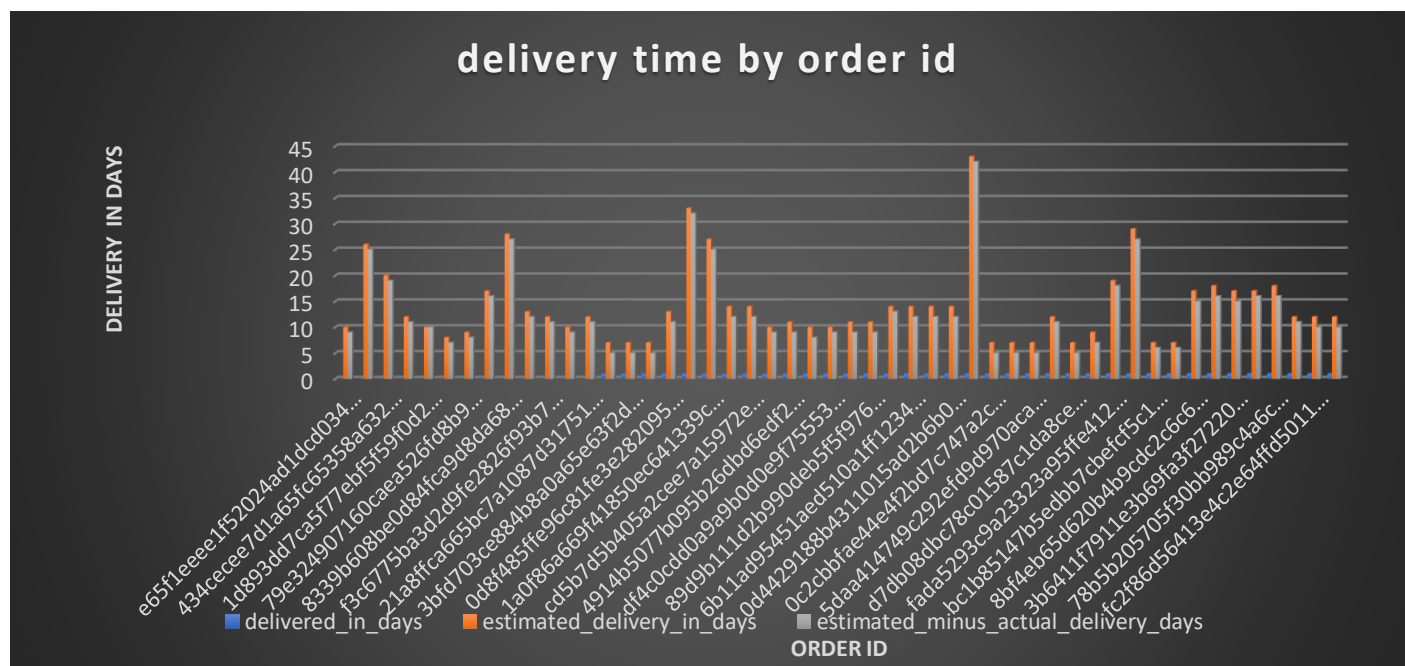
To gain insights into the purchase date, time of delivery and difference between the estimated & actual delivery date OF an order we executed the following query:

## QUERY:

```
SELECT
    order_id,
    DATE_DIFF(order_delivered_customer_date, order_purchase_timestamp, DAY)
    AS delivered_in_days,
    DATE_DIFF(order_estimated_delivery_date, order_purchase_timestamp, DAY)
    AS estimated_delivery_in_days,
    DATE_DIFF(order_estimated_delivery_date, order_delivered_customer_date, DAY)
    AS estimated_minus_actual_delivery_days
FROM
    `target_sql.orders`
WHERE
    DATE_DIFF(order_delivered_customer_date, order_purchase_timestamp, DAY) IS NOT NULL
ORDER BY
    delivered_in_days;
```

## OUTPUT:

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	order_id	delivered_in_days	estimated_delivery_in_days	estimated_minus_actual_delivery_days	
1	e65f1eeee1f52024ad1dcd034...	0	10	9	
2	bb5a519e352b45b714192a02f...	0	26	25	
3	434cecee7d1a65fc65358a632...	0	20	19	
4	d3ca7b82c922817b06e5ca211...	0	12	11	
5	1d893dd7ca5f77ebf5f59f0d20...	0	10	10	
6	d5fbedc85190ba88580d6f82...	0	8	7	
7	79e324907160caea526fd8b94...	0	9	8	
8	38c1e3d4ed6a13cd0cf612d4c...	0	17	16	
9	8339b608be0d84fca9d8da68b...	0	28	27	
10	f349cdb62f69c3fae5c4d7d3f3...	0	13	12	



## INSIGHTS:

- 1) To grow the business we should be well aware of the estimated delivery date and difference between the actual and estimated delivery.
- 2) By analysing the above data we can increase or decrease the number of delivery people so as to meet the estimated delivery date of the order.

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

We estimated the top five states with highest and lowest average freight values in the area Brazil from where the customers purchase.

QUERY:

SELECT

high.customer\_state AS high\_state,  
high.average\_freight\_value AS high\_avg\_freight,  
low.customer\_state AS low\_state,  
low.average\_freight\_value AS low\_avg\_freight

FROM (

SELECT

c.customer\_state,  
ROUND(AVG(p.freight\_value),2) AS average\_freight\_value,  
ROW\_NUMBER() OVER(ORDER BY (ROUND(AVG(p.freight\_value),2))DESC) AS rowval1

FROM

`target\_sql.orders` AS o

JOIN

`target\_sql.order\_items` AS p

ON

o.order\_id = p.order\_id

JOIN

`target\_sql.customers` AS c

ON

o.customer\_id = c.customer\_id

GROUP BY

c.customer\_state

ORDER BY

average\_freight\_value DESC

LIMIT

5 ) AS high

JOIN (

SELECT

c.customer\_state,  
ROUND(AVG(p.freight\_value),2) AS average\_freight\_value,  
ROW\_NUMBER() OVER(ORDER BY (ROUND(AVG(p.freight\_value),2))) AS rowval2

FROM

`target\_sql.orders` AS o

JOIN

`target\_sql.order\_items` AS p

ON

o.order\_id = p.order\_id

JOIN

`target\_sql.customers` AS c

ON

o.customer\_id = c.customer\_id

GROUP BY

c.customer\_state

ORDER BY

average\_freight\_value

LIMIT

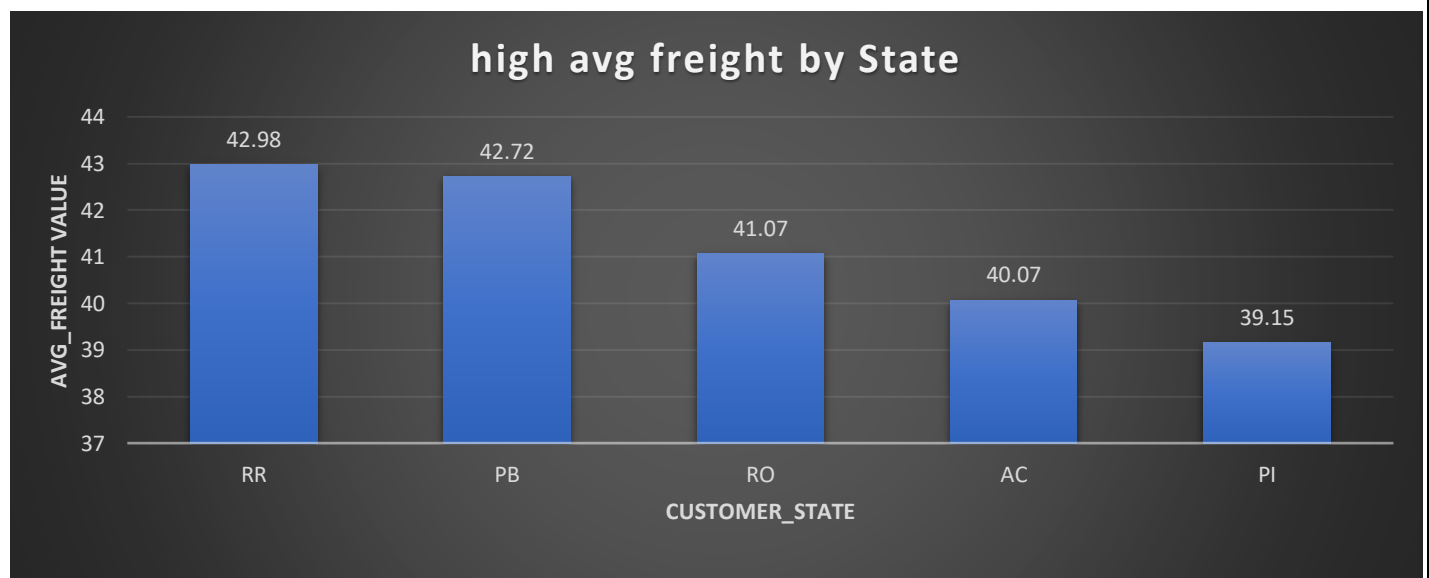
5 ) AS low

ON

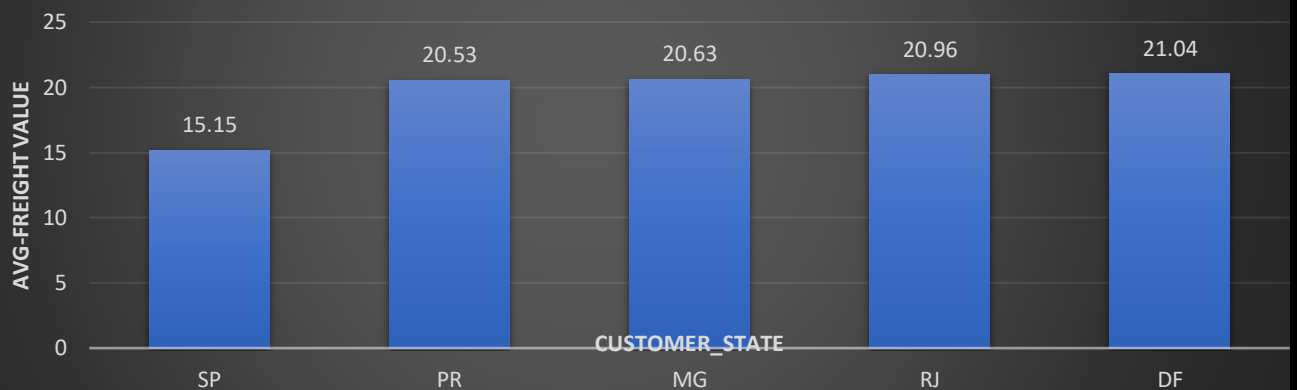
high.rowval1 = low.rowval2;

OUTPUT:

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS		EXECUTION GRAPH
Row	high_state	high_avg_freight	low_state	low_avg_freight		
1	RR	42.98	SP	15.15		
2	PB	42.72	PR	20.53		
3	RO	41.07	MG	20.63		
4	AC	40.07	RJ	20.96		
5	PI	39.15	DF	21.04		



## low avg freight by State



### INSIGHTS:

1)By analysing the above results we found out that the top 5 states which have the highest five states which has the highest average freight value are the following :

-RR  
-PB  
-RO  
-AC  
-PI

2)And the top 5 states which have the lowest five states which has the highest average freight value are the following:

-SP  
-PR  
-MG  
-RJ  
-DF

3)The state with highest average freight value was found to be RR and the state with lowest average freight value was found to be SP.

5.3)Find out the top 5 states with the highest & lowest average delivery time.

To gain insights into the highest and lowest average delivery time the following query was executed:

### QUERY:

WITH

```
cte AS (  
  SELECT  
    c.customer_state,  
    ROUND(AVG(t1.delivery_time),2) AS avg_delivery_time  
  FROM (  
    SELECT  
      *,
```

```

TIMESTAMP_DIFF(order_delivered_customer_date,order_purchase_timestamp,day) AS delivery_time
FROM
    `target_sql.orders` o
WHERE
    order_status='delivered'
    AND order_delivered_customer_date IS NOT NULL
ORDER BY
    order_purchase_timestamp ASC )AS t1
INNER JOIN
    `target_sql.customers` c
ON
    t1.customer_id=c.customer_id
GROUP BY
    c.customer_state
ORDER BY
    avg_delivery_time )
SELECT
    c1.customer_state AS low_state,
    c1.avg_delivery_time AS low_avg_delivery_time,
    c2.customer_state AS high_state,
    c2.avg_delivery_time AS high_avg_delivery_time
FROM (
    SELECT
        *,
        ROW_NUMBER() OVER(ORDER BY cte.avg_delivery_time DESC) AS row_val_2
    FROM
        cte
    ORDER BY
        row_val_2 )AS c2
INNER JOIN (
    SELECT
        *,
        ROW_NUMBER() OVER(ORDER BY cte.avg_delivery_time) AS row_val_1
    FROM
        cte
    ORDER BY
        row_val_1 ) AS c1
ON
    c1.row_val_1=c2.row_val_2
LIMIT
    5;

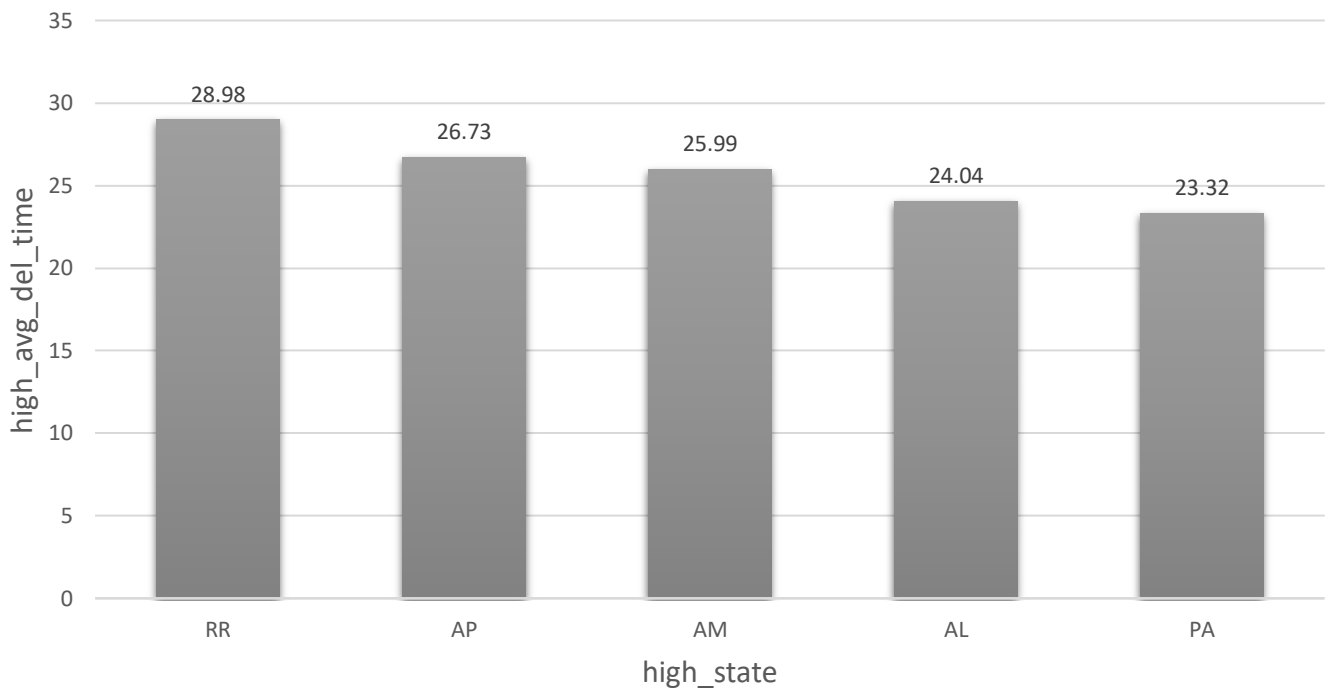
```

OUTPUT

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	low_state ▾	low_avg_delivery_time ▾	high_state ▾	high_avg_delivery_time ▾	
1	SP	8.3	RR	28.98	
2	PR	11.53	AP	26.73	
3	MG	11.54	AM	25.99	
4	DF	12.51	AL	24.04	
5	SC	14.48	PA	23.32	



high avg delivery time by state



#### INSIGHTS:

- 1)from the above data we can observe that the state with the highest average delivery time is Roraima (RR) and the one with the lowest average delivery time is São Paulo (SP).
- 2)After analysing the data we can conclude Roraima (RR) takes lesser time among all the 27 states from where the customers of Brazil shop while São Paulo (SP) takes longest time to deliver the order to the customers.
- 3)Therefore we can say that Roraima has the best delivery strategies and professionals.

5.4)Find out the top 5 state 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.

We have executed the following SQL query in order to find out in which state the order delivery is even before the estimated date of delivery.

#### QUERY:

SELECT

```
c.customer_state,  
ROUND(AVG(DATE_DIFF(order_delivered_customer_date, order_purchase_timestamp, DAY)), 2)  
AS avg_time_to_delivery,  
ROUND(AVG(DATE_DIFF(order_estimated_delivery_date, order_delivered_customer_date, DAY)), 2)  
AS avg_diff_estimated_delivery  
FROM
```



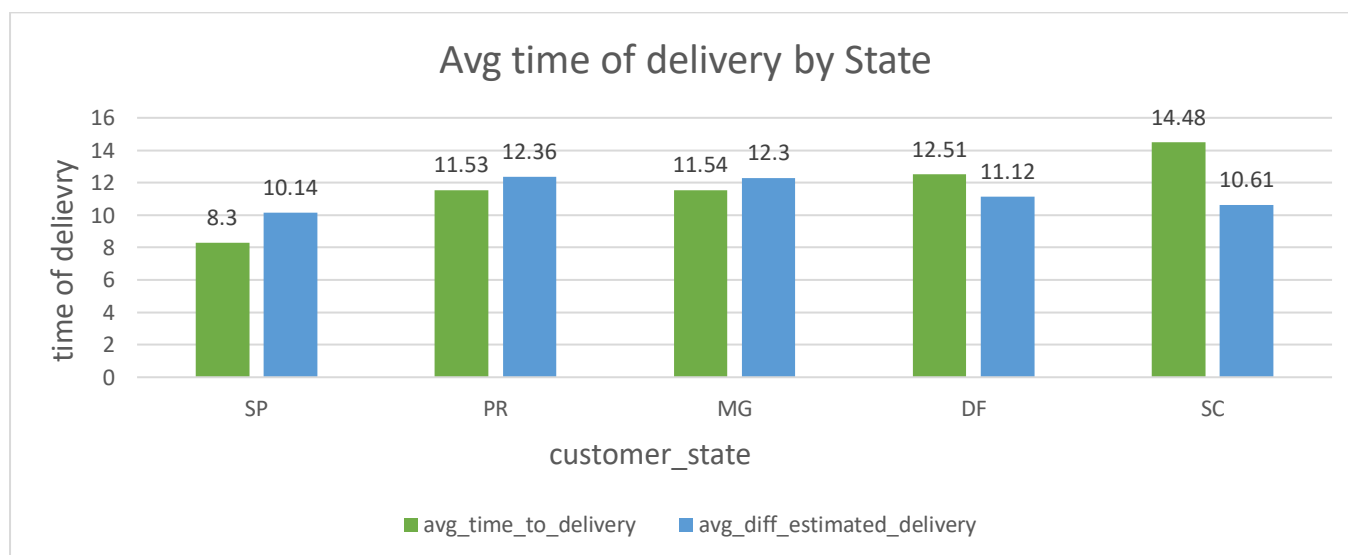
```

`target_sql.orders` o
JOIN
`target_sql.customers` c ON o.customer_id = c.customer_id
WHERE
DATE_DIFF(order_purchase_timestamp, order_delivered_customer_date, DAY) IS NOT NULL
AND
DATE_DIFF(order_estimated_delivery_date, order_delivered_customer_date, DAY) IS NOT NULL
GROUP BY
c.customer_state
ORDER BY
avg_time_to_delivery
LIMIT
5;

```

OUTPUT:

JOB INFORMATION	RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	customer_state	avg_time_to_delivery	avg_diff_estimated_delivery	
1	SP	8.3	10.14	
2	PR	11.53	12.36	
3	MG	11.54	12.3	
4	DF	12.51	11.12	
5	SC	14.48	10.61	



INSIGHTS:

- 1) The above results suggest that the SP, PR, MG, DF and SC are the top 5 states where the order delivery is really fast as compared to the estimated date of delivery.
- 2) Analysing the delivery time is important for businesses to optimize their operations and improve customer satisfaction.

3) Analysing state-wise patterns through SQL queries, companies can gain valuable insights into sales trends, optimize their logistics, and make data-driven decisions to enhance their overall efficiency and customer experience.

Q.6)Analysis based on the payments.

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

To determine the different payment types through which the customers of Brazil pay the bill we analysed the month-on-month no. of orders placed using different payment types.

QUERY:

SELECT

```
p.payment_type,  
extract(month from order_purchase_timestamp) as Month,  
COUNT(o.order_id) as Total_orders
```

FROM

```
`target_sql.orders` o
```

INNER JOIN

```
`target_sql.payments` p
```

ON

```
o.order_id=p.order_id
```

GROUP BY

```
p.payment_type,  
Month
```

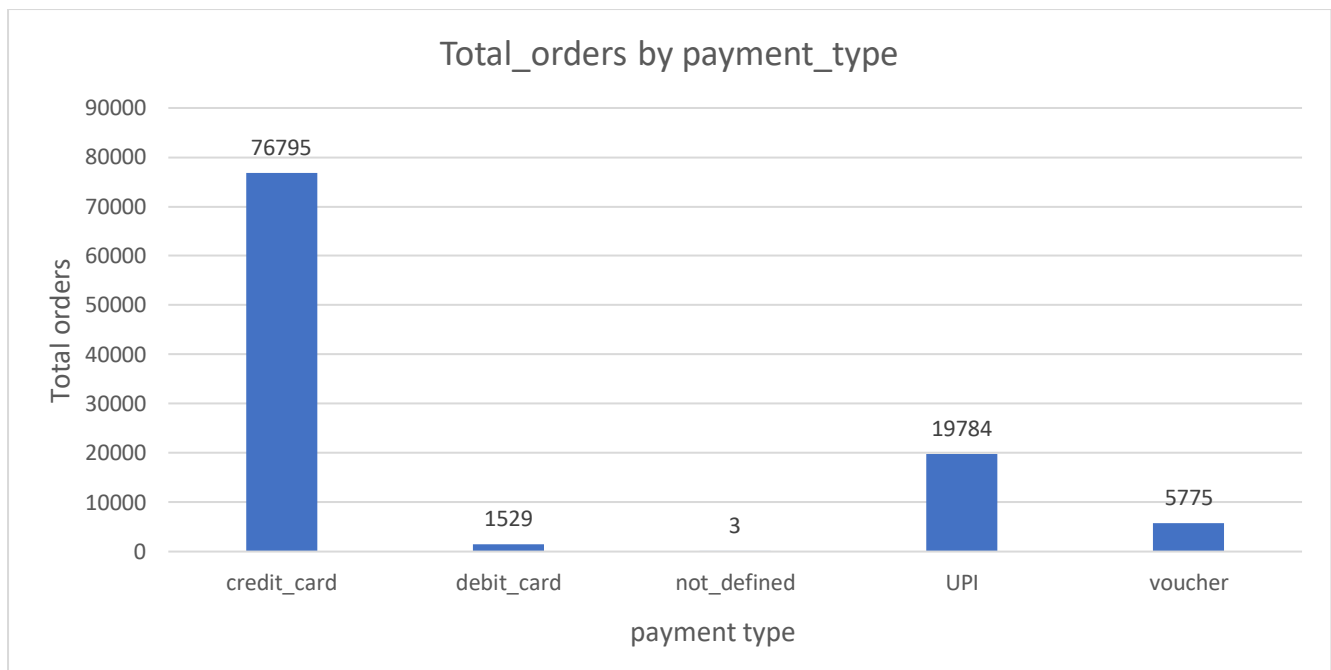
ORDER BY

```
1,
```

```
2;
```

OUTPUT:

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS		EXECUTION GRAPH
Row	payment_type	Month	Total_orders			
1	UPI	1	1715			
2	UPI	2	1723			
3	UPI	3	1942			
4	UPI	4	1783			
5	UPI	5	2035			
6	UPI	6	1807			
7	UPI	7	2074			
8	UPI	8	2077			
9	UPI	9	903			
10	UPI	10	1056			



#### INSIGHTS:

- 1) By analysing the month-on-month no. of orders placed using different payment types we found out that among all the payment types most of the customers pay their bills through credit card.
- 2) On the other hand very a smaller number of people pay through debit card.
- 3) From the above data we can say that credit card is the most preferred payment type among the customers of Target ,it can be due to the benefits they get by paying through credit card.
- 4) Also, we can see that maximum orders are placed mostly between the month of May and August.

6.2) Find the no. of orders placed on the basis of the payment instalments that have been paid.

Finally, we executed a SQL query to find the no. of orders placed on the basis of the payment installments that have been paid.

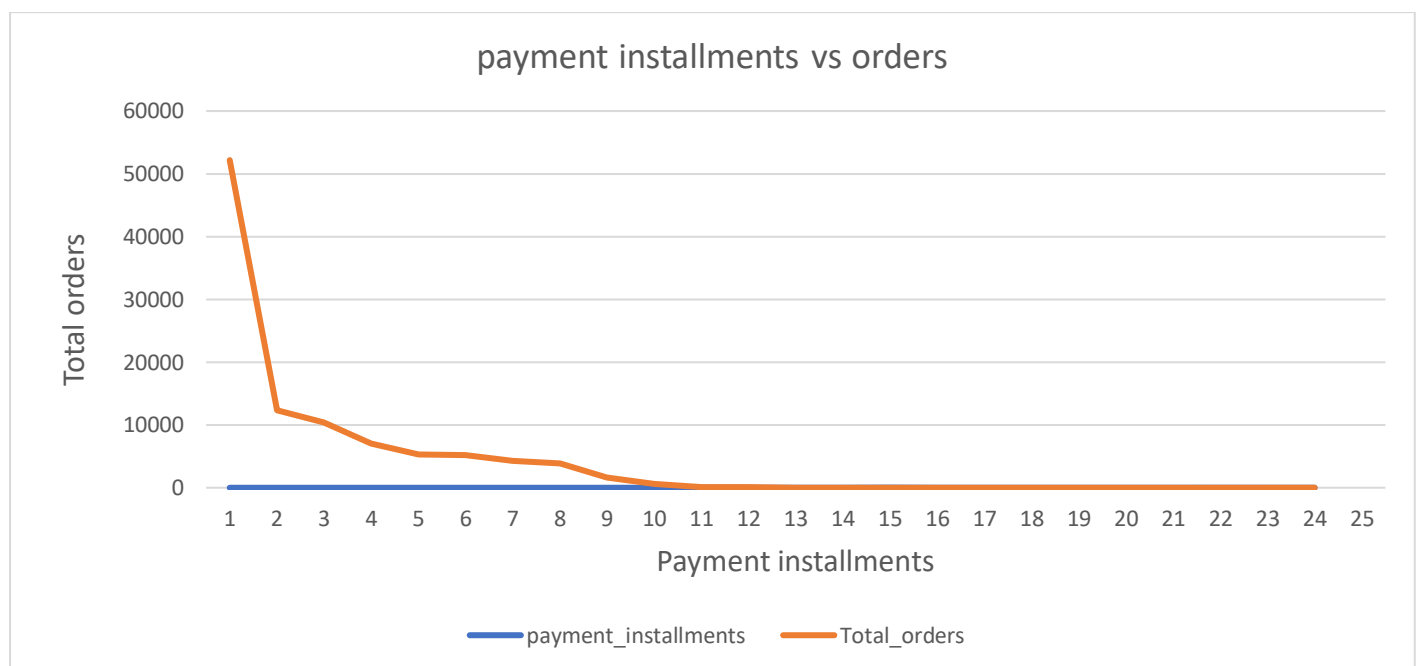
#### QUERY:

```
SELECT
    COUNT(o.order_id) as Total_orders,
    p.payment_installments
FROM
    `target_sql.orders` o
INNER JOIN
    `target_sql.payments` p
ON
    o.order_id=p.order_id
WHERE
    o.order_status != 'canceled'
GROUP BY
```

payment\_installments  
ORDER BY  
Total\_orders DESC;

OUTPUT:

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	payment_installment	Total_orders			
1	1	52184			
2	2	12353			
3	3	10392			
4	4	7056			
5	10	5292			
6	5	5209			
7	8	4239			
8	6	3898			
9	7	1620			
10	9	638			



#### INSIGHTS:

- 1) From the above results we can observe that maximum number of orders that is 52184 which were placed by the customers of Target have only one payment installments and maximum payment installments is 24 which have 24 Orders.
- 2) In order to enhance the customer experience and overall growth of the company the analysis of payment types and payment installments preferences is crucial.
- 3) We can conclude that most of the most of customers prefer single payment installment.
- 4) These insights can help the company to improve their payment strategies.

### INSIGHTS BASED ON THE ANALYSIS

- 1) There was 136.98% increase in the cost of orders from the year 2017 to 2018.
- 2) Mega events like The Olympics Extravaganza(2016), World Cup Frenzy(2018), Carnival celebrations & few more seasonal purchases were the reason for higher sales in certain months.
- 3) Most of the customers of Target shop during afternoon.
- 4) Customers of the Target are mostly from the state São Paulo (SP).
- 5) Maximum orders took place from the month of May to August.
- 6) Maximum number of orders that is 52184 which were placed by the customers of Target have only one payment installments.
- 7) From the analysis we found that most of the customers prefer paying their bills through credit card.
- 8) SP, PR, MG, DF and SC are the top 5 states where the order delivery is really fast as compared to the estimated date of delivery.

### RECOMMENDATIONS BASED ON THE ANALYSIS

- 1) As we can see that most of the customers of Target shop during afternoon followed by night so we should increase the employees during those hours in order to provide smooth customer experience.
- 2) Most of the customers are from São Paulo (SP) so we should make market strategies like collaborating with other sellers from those states, advertising its products, providing offers, discounts during festive seasons so that people from other states too order in high numbers.
- 3) People orders in maximum number during the month of May to August so we should maintain excellent customer support and service during those months and we should promote our business on social media networks throughout the year, we should collaborate with influencers so that we get more customers throughout the year.
- 4) Improve shipping processes so that customers get the products as soon as it can be.
- 5) Improve in technology like keeping the website fresh, offering personalized recommendations to customers so as to improve e-commerce experience.
- 6) Keep an eye on other competitive companies so that we can make strategies accordingly in order to improve the business.
- 7) Evaluate pricing and freight fees to ensure competitiveness in the market while maximizing revenue and profitability. Consider increasing prices or adjusting freight fees as appropriate.
- 8) We should apply discounts and offers during off peak periods to gain more customers.