

Business Case - Scaler - Feb '24 batch

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

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

Ans.

SELECT

COLUMN_NAME, DATA_TYPE

FROM

`scaler-dsml-sql-410603.Business_Case_Target_SQL.INFORMATION_SCHEMA.COLUMNS`

WHERE TABLE_NAME = 'customers_csv';

Screenshot

Query results			SAVE RESULTS	EXPLORE DATA	
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

- There are two columns stating the same data which might not be the same but having both in the same table is confusing. Those columns are 'customer_id' and 'customer_unique_id' .
- There are 4 columns with text/string data type and only 1 column with integer data type.

Recommendation

- Two columns ideally conveying the same thing differently e.g., 'customer_id' and 'customer_unique_id' should not be there in the table
- Get the time range between which the orders were placed.

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

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

Ans.

select

min(order_purchase_timestamp) as `1st order time` ,

max(order_purchase_timestamp) as `last order time`

from

`Business_Case_Target_SQL.orders`

Screenshot

Query results		SAVE RESULTS		EXPLORE DATA	
JOB INFORMATION	RESULTS	CHART	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	1st order time	last order time			
1	2016-09-04 21:15:19 UTC	2018-10-17 17:30:18 UTC			

Insights

- Very 1st order was placed on 4th Sep 2024 at 3:19 pm UTC,
- Last order was placed on 17th Oct 2018 at 5:30 pm UTC,
- Rest of all orders were placed in the time range between (1. And 2.)
- Both 1st and last order are done at noon time.
- Most important time zone here is UTC which means 'Universal Time Coordinated' - it is a reference time zone for adding and subtracting hours to represent any other country's time zone.

Recommendation

- Time zone can be dedicated to any country e.g., UKT for UK , IST for India. Else no other recommendations for now.

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

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

Ans.

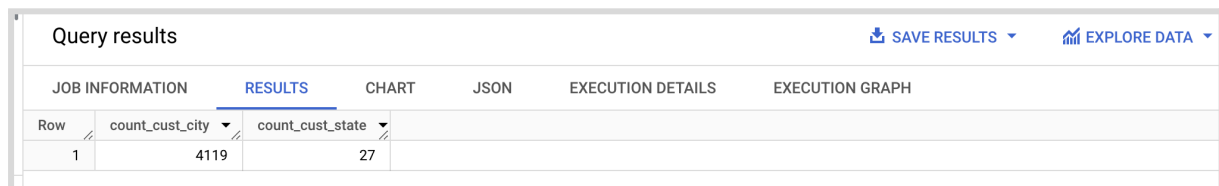
select

```
count(distinct customer_city) `count_cust_city` ,  
count(distinct customer_state) `count_cust_state`
```

from

```
`Business_Case_Target_SQL.customers_csv` ;
```

Screenshot



Query results				SAVE RESULTS		EXPLORE DATA	
JOB INFORMATION		RESULTS		CHART		JSON	
EXECUTION DETAILS		EXECUTION GRAPH					
Row	count_cust_city	count_cust_state					
1	4119	27					

Insights

a. There are distinct 4119 cities and distinct 27 states from which customers ordered during 2016 - 2018

Recommendation

a. Based on this data cluster of order done from specific city and state can be figured out. Also, sales for different product in specific area can be studied and overall sales can be improved in active areas. More areas can be added by introducing products there.

Q-2. In-depth Exploration:

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

Ans.

SELECT

EXTRACT(YEAR FROM order_purchase_timestamp) AS `Order_year`,

EXTRACT(MONTH FROM order_purchase_timestamp) AS `Order_month`,

CONCAT(EXTRACT(MONTH FROM order_purchase_timestamp), '/', EXTRACT(YEAR FROM order_purchase_timestamp)) AS `Order_month_year`,

COUNT(order_id) AS `Count_orders`

FROM

`Business_Case_Target_SQL`.`orders`

GROUP BY

`Order_year`, `Order_month`, `Order_month_year`

ORDER BY

`Order_year` ASC, `Order_month` ASC;

Screenshot

Row	Order_year	Order_month	Order_month_year	Count_orders
1	2016	9	9/2016	4
2	2016	10	10/2016	324
3	2016	12	12/2016	1
4	2017	1	1/2017	800
5	2017	2	2/2017	1780
6	2017	3	3/2017	2682
7	2017	4	4/2017	2404
8	2017	5	5/2017	3700
9	2017	6	6/2017	3245
10	2017	7	7/2017	4026

Insights

a). Based on the query 25 rows are being generated but here we can paste only the top ten rows as per the rules. As per top 10 rows it is observed that no. of orders are increased to 1000 times from 09/2016 to 07/2017,

- b). Based on all 25 rows analysis from 09/2016 to 10/2018 orders have increased to 1500 times until 08/2018 (increase mentioned is over the whole period of time in a linear way). But suddenly orders are seen in single digit for 09/2018 and 10/2018.
- c). There can be many reasons for the sudden down trend in orders in 2018.

Recommendation

- a). Feedback for down trend in orders for 09/2018 and 10/2018 to be done to get back the order trends as in previous months,
- b). With each passing month as observed in top 10 rows orders are getting doubled -> reason can be found out and still more orders can be received,
- c). Out of order placed -> cancelled and completed orders should be segregated and specific reasons for cancelled orders should be figured out.
- d). Based on the season, what is the order trend ? by optimal analysis can increase no. of orders placed.

Q-2. In-depth Exploration:

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

Ans.

SELECT

EXTRACT(YEAR FROM order_purchase_timestamp) AS `Order_year`,

EXTRACT(MONTH FROM order_purchase_timestamp) AS `Order_month`,

CONCAT(EXTRACT(MONTH FROM order_purchase_timestamp), '/', EXTRACT(YEAR FROM order_purchase_timestamp)) AS `Order_month_year`,

COUNT(order_id) AS `Count_orders`,

FROM

`Business_Case_Target_SQL`.`orders`

GROUP BY

`Order_year`, `Order_month`, `Order_month_year`

ORDER BY

`Order_month`, `Order_year`;

Screenshot

Row	Order_year	Order_month	Order_month_year	Count_orders
1	2017	1	1/2017	800
2	2018	1	1/2018	7269
3	2017	2	2/2017	1780
4	2018	2	2/2018	6728
5	2017	3	3/2017	2682
6	2018	3	3/2018	7211
7	2017	4	4/2017	2404
8	2018	4	4/2018	6939
9	2017	5	5/2017	3700
10	2018	5	5/2018	6873

Insights

- Total no. of rows generated are 25- here as per limit only 10 rows are pasted,
- Max no. of order = 7544 in 11/2017 -> may be pre- Christmas sales,
- 2nd highest no. of orders = 01/2018 -> may be new year shopping,
- Minimum no. of orders or orders in single digit were observed during last months of year 2016 and 2018 -> may be due to beginning and ending of Target company or some technical issues or global problem.,

- e. Other important insight can be observed that as years are passing orders in comparison to months of previous years have increased dramatically -> it may be due to 'company' being famous with each passing year.

Recommendation

- a. Minimum no. of orders or orders in single digit were observed during last months of year 2016 and 2018 -> may be due to beginning and ending of Target company or some technical issues or global problem -> but this insight can be researched and in future "receiving such less no. of orders" situation can be avoided,
- b. Why with each passing year orders have increased dramatically ? and the same strategies should be followed for further growth of companies in upcoming years.

Q-2. In-depth Exploration:

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

- i. 0-6 hrs : Dawn
- ii. 7-12 hrs : Mornings
- iii. 13-18 hrs : Afternoon
- iv. 19-23 hrs : Night

Ans.

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'

ELSE 'Night'

END AS `time_classification`,

COUNT(*) AS `Count_orders`

FROM

`Business_Case_Target_SQL.orders`

GROUP BY

time_classification

ORDER BY

Count_orders DESC;

Screenshot

Row	time_classification ▼	Count_orders ▼
1	Afternoon	38135
2	Night	28331
3	Morning	27733
4	Dawn	5242

Insights

- a. Based on the results in the screen shot it can be concluded that maximum numbers of orders are being placed during 'Afternoon' time. Afternoon orders are more than 1/3rd of total orders,
- b. 2nd maximum orders are being placed at 'Night' time which are nearing to 1/3rd of total orders,
- c. 3rd maximum numbers of orders are placed in morning time and number of orders are approx same with 'count of orders' during night time
- d. Minimum orders are placed during Dawn time which are less than 1/20 th of total orders.

Recommendations

- a. Offers can be rolled out for 'Dawn' time so that people start ordering in 'Dawn' time,
- b. Capacity of site must be increased for afternoon and night time - as per case study of Big Basket the site of big basket got crashed when in afternoon maximum orders were getting places in corona times -> hence for 1 month time the site was quite slow and unavailable for orders -> leading to less profit than what it could have been.
- c. Site related maintenance can be done when there is less traffic on site.
- d. Here we are only seeing the numbers for order placed. Based 'order status' data should be for all the orders except for 'cancelled' and 'rejected' orders.

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

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

Ans.

SELECT

C.customer_state,

FORMAT_TIMESTAMP('%m/%Y', O.order_purchase_timestamp) AS 'MM_YYYY' ,

COUNT(O.order_id) AS order_count

FROM `Business_Case_Target_SQL.customers_csv` AS C

INNER JOIN

`Business_Case_Target_SQL.orders` AS O

ON

C.customer_id = O.customer_id

GROUP BY

C.customer_state, MM_YYYY

Order by

order_count DESC;

Screenshot

Query results					SAVE RESULTS	EX
JOB INFORMATION		RESULTS	CHART	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	customer_state	MM_YYYY	order_count			
1	SP	08/2018	3253			
2	SP	05/2018	3207			
3	SP	04/2018	3059			
4	SP	01/2018	3052			
5	SP	03/2018	3037			
6	SP	11/2017	3012			
7	SP	07/2018	2777			
8	SP	06/2018	2773			
9	SP	02/2018	2703			
10	SP	12/2017	2357			

Results per page: 50 1 – 50 of 565

Insights

- Maximum no. of orders were placed from 'SP' state in the year 2018 in the month of August,
- There is data for all the states based on MM/YYYY format giving a result of 565 rows - so detailed analyses can be done for any area as and when required.
- States like 'SC' , 'PI' , 'RJ' , 'PB' are having orders for a few months in single digit.

Recommendations

- a. Based on data extracted, a question for why some states are having maximum count of orders and why many states are having count of orders as single digit should be figured out.
- b. After figuring out the reason, optimal solutions should be planned and executed and the count of orders should be improved.

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

2).How are the customers distributed across all the states?

Ans.

```
SELECT customer_state,  
       COUNT(DISTINCT customer_unique_id) AS `COUNT_CUST`,  
FROM `Business_Case_Target_SQL.customers_csv`  
GROUP BY customer_state  
ORDER BY COUNT_CUST DESC;
```

Screenshot

Row	customer_state	COUNT_CUST
1	SP	40302
2	RJ	12384
3	MG	11259
4	RS	5277
5	PR	4882
6	SC	3534
7	BA	3277
8	DF	2075
9	ES	1964
10	GO	1952

Results per page: 50 1 – 27 of 27

Insights

a. Total uniques customers as per col “customer_unique_id” are 96,096 out of which maximum no. of unique customers are from SP state followed by other states as seen in screenshot,

b. Minimum no. of customers are from “RR” , “AP” and “AC” -> i.e., having less than 100 no. of unique customers.

Recommendations

a. Question of why there are less than 100 unique customer in few states should be answered and henceforth no. of customers should be increased after analysing the data.

b. Strategy for states with maximum customers should be improved to increase no. of customers in those states in future.

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

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

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

Ans.

```
WITH yearly_costs AS (  
    SELECT  
        EXTRACT(YEAR FROM o.order_purchase_timestamp) AS Year,  
        ROUND(SUM(p.payment_value), 0) AS Total_cost  
    FROM  
        `Business_Case_Target_SQL.payments` AS p  
    INNER JOIN  
        `Business_Case_Target_SQL.orders` AS o  
    ON  
        p.order_id = o.order_id  
    WHERE  
        EXTRACT(MONTH FROM o.order_purchase_timestamp) BETWEEN 1 AND 8  
    GROUP BY  
        Year  
)  
SELECT  
    Year,  
    Total_cost,  
    ROUND((Total_cost - LAG(Total_cost) OVER (ORDER BY Year)) / LAG(Total_cost) OVER  
(ORDER BY Year) * 100, 2) AS Percent_increase  
FROM  
    yearly_costs  
ORDER BY  
    Year;
```

Screenshot

Query results				
JOB INFORMATION		RESULTS	CHART	JSON
Row	Year ▼	Total_cost ▼	Percent_increase ▼	
1	2017	3669022.0	null	
2	2018	8694734.0	136.98	

Insights

a. By looking at data in table one can conclude that inflation might have increased by 136% from Year 2017 to 2018 but that would be false statement as here sum of cost is more than double the cost in 2017 due to increase in no.of orders

Recommendations

- To get exact understanding on % increase in inflation head-to-head comparison with cost of product with same id in 2017 to 2018 should be done,
- Comparison can be done product category wise to understand inflation category wise.

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

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

Ans.

```
SELECT C.customer_state,
       ROUND(AVG(I.price),0) as `AVG_price`,
       ROUND(SUM(I.price),0) as `TOTAL_price`,
FROM `Business_Case_Target_SQL.order_items` as I
LEFT JOIN
`Business_Case_Target_SQL.orders` as O
ON I.order_id = O.order_id
LEFT JOIN
`Business_Case_Target_SQL.customers_csv` as C
ON O.customer_id = C.customer_id
Group BY C.customer_state
ORDER BY TOTAL_Price DESC, AVG_price DESC;
```

Screenshot

JOB INFORMATION		RESULTS	CHART	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	customer_state ▼	AVG_price ▼	TOTAL_price ▼			
1	SP	110.0	5202955.0			
2	RJ	125.0	1824093.0			
3	MG	121.0	1585308.0			
4	RS	120.0	750304.0			
5	PR	119.0	683084.0			
6	SC	125.0	520553.0			
7	BA	135.0	511350.0			
8	DF	126.0	302604.0			
9	GO	126.0	294592.0			
10	ES	122.0	275037.0			
Results per page: 50 ▼ 1 – 27 of 27						

Insights

- Maximum sales has happened in state of "SP" as sum(total price) figure is seen maximum for the same,
- But average price for max of total is lesser may be due to more items of less cost were purchased from SP state,

- c. Data for 'Total cost' and 'Avg cost' can be observed for each individual state for the time period from 2016-2018.

Recommendations

- a. Proper analyses of the data can be used to improve the sales for each states and figures can be increased.

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

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

Ans.

```
SELECT C.customer_state,
       ROUND(AVG(I.freight_value),0) as `AVG_freight_value`,
       ROUND(SUM(I.freight_value),0) as `TOTAL_freight_value`,
FROM `Business_Case_Target_SQL.order_items` as I
LEFT JOIN
`Business_Case_Target_SQL.orders` as O
ON I.order_id =O.order_id
LEFT JOIN
`Business_Case_Target_SQL.customers_csv` as C
ON O.customer_id = C.customer_id
Group BY C.customer_state
ORDER BY TOTAL_Freight_value DESC, AVG_freight_value DESC;
```

Screenshot

JOB INFORMATION		RESULTS	CHART	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	customer_state ▾	AVG_freight_value ▾	TOTAL_freight_value ▾			
1	SP	15.0	718723.0			
2	RJ	21.0	305589.0			
3	MG	21.0	270853.0			
4	RS	22.0	135523.0			
5	PR	21.0	117852.0			
6	BA	26.0	100157.0			
7	SC	21.0	89660.0			
8	PE	33.0	59450.0			
9	GO	23.0	53115.0			
10	DF	21.0	50625.0			
Results per page: 50 ▾ 1 – 27 of 27						

Insights

a. State SP is having maximum 'total of freight value' but at the same time SP is not having maximum of 'average freight value' ,

b. State RR has a minimum of 'total of freight value' but not minimum of average freight value.

Recommendations

a. How efficiently freight can be used should be determined from the output of result data in order to make savings on operating costs.

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

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:

- $\text{time_to_deliver} = \text{order_delivered_customer_date} - \text{order_purchase_timestamp}$
- $\text{diff_estimated_delivery} = \text{order_delivered_customer_date} - \text{order_estimated_delivery_date}$

Ans.

SELECT *

FROM `Business_Case_Target_SQL.orders`

SELECT DATE_DIFF(order_delivered_customer_date, order_purchase_timestamp, DAY) AS
`time_to_deliver_days` ,

DATE_DIFF(order_delivered_customer_date, order_estimated_delivery_date, DAY) AS
`diff_estimated_delivery_days`

FROM `Business_Case_Target_SQL.orders`

Screenshot

Row	time_to_deliver_days	diff_estimated_delivery_days
1	30	12
2	30	-28
3	35	-16
4	30	-1
5	32	0
6	29	-1
7	43	4
8	40	4
9	37	1
10	33	5

Results per page: 50 1 – 50 of 99441

Insights

a. We can observe that time_to deliver in days is positive and will always be as it is the difference between delivery done in no. of days after order is placed,

b. Also in column-2 we can see that there are negative figures which indicate that orders were delivered much earlier than the estimated delivery date which is a good sign for business.

Positive figures indicate the delay which should be worked out.

Recommendations

a. But looking at delivery days of amazon in India Target should take less time to deliver - maybe not more than a week. So that more orders can be received if delivery time is decreased.

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

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

Ans.

Query for top 5 states with highest & lowest average freight value changes only in order by clause and derived results can be seen in screenshots pasted.



```
SELECT C.customer_state,
       ROUND(AVG(I.freight_value),0) as `AVG_freight_value`,
       ROUND(SUM(I.freight_value),0) as `TOTAL_freight_value`,
FROM `Business_Case_Target_SQL.order_items` as I
LEFT JOIN
`Business_Case_Target_SQL.orders` as O
ON I.order_id = O.order_id
LEFT JOIN
`Business_Case_Target_SQL.customers_csv` as C
ON O.customer_id = C.customer_id
Group BY C.customer_state
ORDER BY AVG_freight_value DESC;
```

Screenshot

a. Top 5 states with highest average of freight value

Query results					SAVE RESULTS	E
JOB INFORMATION		RESULTS	CHART	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	customer_state ▼	AVG_freight_value	TOTAL_freight_value			
1	PB	43.0	25720.0			
2	RR	43.0	2235.0			
3	RO	41.0	11417.0			
4	AC	40.0	3687.0			
5	PI	39.0	21218.0			
					Results per page: 50 ▼	1 – 27 of 27

b. Top 5 states with lowest freight value

Query results					 SAVE RESULTS ▾	 E
JOB INFORMATION		RESULTS	CHART	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	customer_state ▾	AVG_freight_value ▾ ↑	TOTAL_freight_value			
1	SP	15.0	718723.0			
2	RJ	21.0	305589.0			
3	PR	21.0	117852.0			
4	SC	21.0	89660.0			
5	DF	21.0	50625.0			
					Results per page: 50 ▾	1 – 27 of 27

Insights

- a. State PB is having highest freight value of 43,
- b. State SP has the lowest freight value of 15.

Recommendations

- a. Why some states are having highest and lowest average freight value is to be figured out and operating costs should be decreased by doing some strategically reduction in freight value.

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

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

Ans.

Here, order by clause will be reversed by ascending to descending to find top 5 states with highest and lowest average delivery time.

```
SELECT C.customer_state,  
       Round(AVG(DATE_DIFF(O.order_delivered_customer_date, O.order_purchase_timestamp,  
DAY)),1) AS `avg_time_to_deliver_days`  
FROM `Business_Case_Target_SQL.orders` as O  
LEFT JOIN  
`Business_Case_Target_SQL.customers_csv` as C  
ON O.customer_id = C.customer_id  
GROUP BY C.customer_state  
ORDER BY avg_time_to_deliver_days DESC;
```

Screenshot

a. Top 5 states with highest average delivery time

Query results			SAVE RESULTS	
JOB INFORMATION			RESULTS	CHART
			JSON	EXECUTION DETAILS
			EXECUTION GRAPH	
Row	customer_state	avg_time_to_deliver		
1	RR	29.0		
2	AP	26.7		
3	AM	26.0		
4	AL	24.0		
5	PA	23.3		
			Results per page:	50
				1 – 27 of 27

b. Top 5 states with lowest delivery time

Query results

SAVE RESULTS

JOB INFORMATIONRESULTSCHARTJSONEXECUTION DETAILSEXECUTION GRAPH

Row	customer_state	avg_time_to_deliver_days
1	SP	8.3
2	PR	11.5
3	MG	11.5
4	DF	12.5
5	SC	14.5

Results per page: 501 – 27 of 27

Insights

a. SP is taking very less delivery time and hence results in prior questions we observe that there are maximum orders and cost from SP state and lowest average freight value too.

Meaning lesser delivery time leads to more no. of orders and average freight value also goes down,

b. Also at the same time in state RR delivery time is highest and looking at the results in prior questions RR state is having 2nd highest freight value and is nowhere in top 10 list of highest total cost hence it is receiving much less orders due to it's longer delivery time

Recommendations

a. For states taking higher delivery time - some workout should be done to lessen the delivery time leading to more no. of orders leading to less avg-freight that means saving on operating cost and increase in revenue and margin.

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

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.

Ans.

States with higher negative values(as seen in screenshot) have delivered faster than the estimated time of delivery.

```
SELECT C.customer_state,
       ROUND(AVG(DATE_DIFF(order_delivered_customer_date,
order_estimated_delivery_date, DAY)),1) AS `AVG_diff_estimated_delivery_days`
FROM `Business_Case_Target_SQL.orders` as O
LEFT JOIN
`Business_Case_Target_SQL.customers_csv` as C
ON O.customer_id = C.customer_id
WHERE O.order_status = 'delivered'
GROUP BY C.customer_state
ORDER BY AVG_diff_estimated_delivery_days ASC;
```

Screenshot

Query results			SAVE RESULTS	E
JOB INFORMATION			RESULTS	CHART
			JSON	EXECUTION DETAILS
			EXECUTION GRAPH	
Row	customer_state ▼	AVG_diff_estimated		
1	AC	-19.8		
2	RO	-19.1		
3	AP	-18.7		
4	AM	-18.6		
5	RR	-16.4		
			Results per page: 50 ▼	1 – 27 of 27

Insights

a. States AC to RR has delivered orders in advance of 15-20 days prior to estimated delivery time.

Recommendations

- a. Based on insight one can recommend that when a customer sees more days in estimated delivery time it sometimes becomes obvious that customer would go to a nearby shop and buy the product to avoid delay - which will get Target lesser orders.
- b. It is highly recommended to estimate optimal delivery time to get more orders as top states with higher total cost are not seen in the screenshot with higher estimated delivery time.

Q-6. Analysis based on the payments:


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


Ans.

```
SELECT P.payment_type,  
       count(O.order_id) as `Orders_count`,  
       FORMAT_TIMESTAMP('%m/%Y', order_purchase_timestamp) AS `month_year`  
FROM `Business_Case_Target_SQL.orders` as O  
INNER JOIN  
`Business_Case_Target_SQL.payments` as P  
ON O.order_id = P.order_id  
GROUP BY P.payment_type, FORMAT_TIMESTAMP('%m/%Y', order_purchase_timestamp)  
ORDER BY `Orders_count` DESC;
```

Screenshot

Query results

 SAVE RESULTS ▾



JOB INFORMATION

RESULTS

CHART

JSON

EXECUTION DETAILS

EXECUTION GRAPH

Row	payment_type ▾	Orders_count ▾	month_year ▾	
1	credit_card	5897	11/2017	
2	credit_card	5691	03/2018	
3	credit_card	5520	01/2018	
4	credit_card	5497	05/2018	
5	credit_card	5455	04/2018	
6	credit_card	5253	02/2018	
7	credit_card	4985	08/2018	
8	credit_card	4813	06/2018	
9	credit_card	4755	07/2018	
10	credit_card	4377	12/2017	

Results per page: 50 ▾ 1 – 50 of 90

Insights

- As seen in the screenshot above most of the payments are done through credit card.
- Payment of more than 75% of orders are done through credit card.

Recommendations

- Based on which bank credit cards are used at most Target should make a deal with the bank that if 25% or more payments are done through any particular bank then the bank should offer a special discount to customers as well as Target. This will help to fetch more orders and reduce the operating cost or increase the operating profit.

b. Based on minimum payments through different other payment types Target should make a deal that if you want your platform to be used for payments offer Target and consumers with attractive discounts.

c. With recommendation - a and b Target will gain more orders and plus on operating profit.

Q-6. Analysis based on the payments:

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

Ans.

```
SELECT COUNT(order_id) AS `NUMBER_ORDERS` ,  
       payment_installments  
FROM `Business_Case_Target_SQL.payments`  
WHERE payment_sequential >=1  
GROUP BY payment_installments  
ORDER BY NUMBER_ORDERS DESC;
```

Screenshot

Query results				SAVE RESULTS ▾		📊	
JOB INFORMATION		RESULTS	CHART	JSON	EXECUTION DETAILS	EXECUTION GRAPH	
Row	NUMBER_ORDERS	payment_installments					
1	52546	1					
2	12413	2					
3	10461	3					
4	7098	4					
5	5328	10					
6	5239	5					
7	4268	8					
8	3920	6					
9	1626	7					
10	644	9					
				Results per page: 50 ▾		1 – 24 of 24	

Insights

- Orders with no. of instalment =1 are more than 50% of total orders,
- There are comparatively less orders with higher no. of payment instalments.

Recommendations

- Customers are likely to make payment in 1 instalment but this can be true for specific categories of product.
- Need to get more data on product and product category where customers are likely to make payments in instalments.