TARGET SQL PROJECT

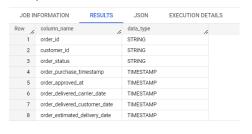
Apr 18, 2023

1)

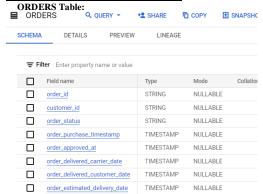
[a]

SELECT column_name,data_type from sqllproject.INFORMATION_SCHEMA.COLUMNS WHERE TABLE_NAME='ORDERS'

Query results



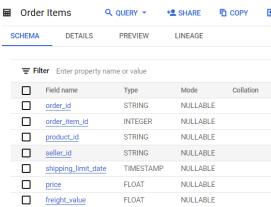
DATA TYPES OF COLUMNS IN A TABLE:



CONCLUSION:

The Data Type of columns, order_id is STRING, customer_id is STRING, order_status is STRING, order_purchases_timestamp is TIMESTAMP, order_approved_at is TIMESTAMP, order_delivered_carrier_date is TIMESTAMP, order_delivered_customer_date is TIMESTAMP, order_estimate_delivery_date is TIMESTAMP.

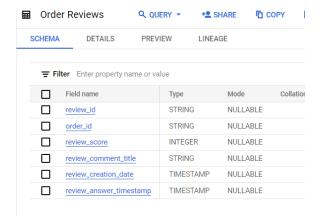
ORDER ITEMS Table :



CONCLUSION:

The Data Type of columns, order_id is STRING, order_item_id is INTEGER, productid is STRING, seller_id is STRING, shipping_limit_date is TIMESTAMP, price is FLOAT, freight_value is FLOAT.

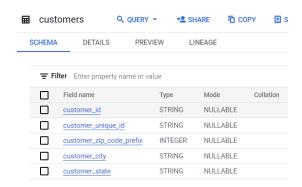
ORDER REVIEWS Table :



CONCLUSION:

The Data Type of the columns, review_id is STRING, order_id is STRING, review_score is INTEGER, review_comment_titleis STRING, review_creation_date is TIMESTAMP, review_answer_timestamp is TIMESTAMP.

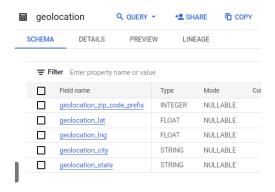
CUSTOMERS Table:



CONCLUSION:

 $The \ Data \ Type \ of the \ columns \ \ customer_id \ is \ \textit{STRING}, \ customer_unique_id \ is \ \textit{STRING}, \ customer_zip_code_prefix \ is \ \textit{INTEGER}, \ description \ and \ description \ desc$ customer_id is STRING, customer_state is STRING.

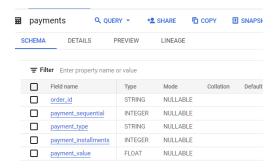
GEOLOCATION Table:



CONCLUSION:

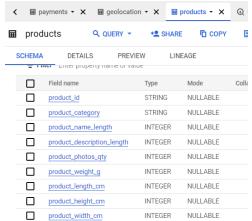
The Data Type of the columns geolocation_zip_code_prefix is an *INTEGER*, geolocation_lat is *FLOAT*, geolocation_lng is *FLOAT*, geolocation_city is *STRING*, geolocation_state is *STRING*.

PAYMENTS Table:



The Data Type of the columns order_id is STRING, payment_sequential is an INTEGER, payment_type is STRING, payment_installments is an INTEGER, payment_value is ELOAT

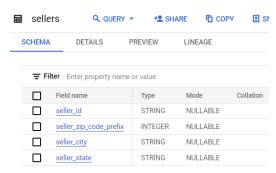




COMCLUSION:

The Data Types of columns product_id is STRING, product_category is STRING, product_name_length is an INTEGER, product_description_length is an INTEGER, product_hotos_qty is an INTEGER, product_weight_g is an INTEGER, product_height_cm is an INTEGER, product_height_cm is an INTEGER.

SELLERS Table:



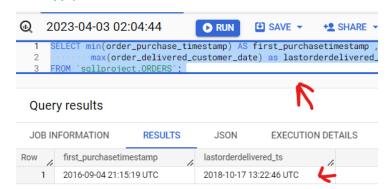
CONCLUSION:

The data type of the column seller_id is STRING, seller_zip_code_prefix is INTEGER, seller_city is STRING, seller_state is STRING.

[b] Time period for which data is given is:

SELECT min(order_purchase_timestamp) AS first_purchasetimestamp,
max(order_delivered_customer_date) as lastorderdelivered_ts

FROM `sqllproject.ORDERS`;



CONCLUSION:

Time period is technically in between $\emph{first_purchase}$ and $\emph{lastorderdelevered_ts}$.

What is meant by that is frankly speaking time period is in between first time purchase order took place and the last order that was delivered in the first case, in the second case we have taken the max and min of all the time related condition.

[c] CITIES and STATES of customers during the given period:

SELECT DISTINCT cust.customer_city, cust.customer_state

FROM 'sqllproject.customers' cust

JOIN `sqllproject.ORDERS` o ON cust.customer_id = o.customer_id

WHERE o.order_delivered_customer_date BETWEEN order_purchase_timestamp and order_estimated_delivery_date

or

SELECT DISTINCT customer_city, customer_state

FROM `sqllproject.customers`

is also more than sufficient.

JOB IN	JOB INFORMATION		JSON	EXECUTION DETAILS
Row /	customer_city	6	customer_state	
1	acu		RN	
2	acu		RN	
3	acu		RN	
4	ico		CE	
5	ico		CE	
6	ico		CE	
7	ico		CE	
8	ico		CE	
9	ico		CE	
4.0			0.5	

THESE are few customer cities and states.

```
(2)
[a] To find the growing trend in e commerce query is given below:

SELECT extract(YEAR from ord.order_purchase_timestamp) as Year, count(distinct ord.order_id) as ORDERSS,sum(ordi.price) as TOT_PRICE,sum(pm.payment_value) as TPT_PAY
FROM

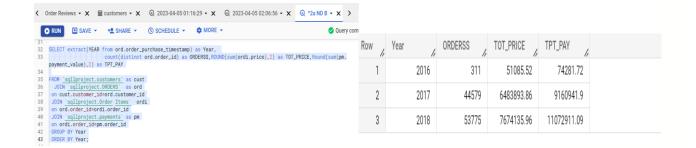
'sqllproject.ORDERS' as ord

JOIN 'sqllproject.Order Items' ordi
on ord.order_id=ordi.order_id
```

on ordi.order_id=pm.order_id GROUP BY Year

JOIN 'sqllproject.payments' as pm

Order By Year



CONCLUSION: The fact that you can see the increase in distinct orders from the column orderss proves that there is a growing trend on

[B] SELECT

CASE

WHEN EXTRACT(HOUR FROM ord.order_purchase_timestamp)>=5 AND EXTRACT(HOUR FROM ord.order_purchase_timestamp)<9

THEN 'DAWN'
WHEN EXTRACT(HOUR FROM ord.order_purchase_timestamp) >= 9 AND EXTRACT(HOUR FROM ord.order_purchase_timestamp) < 12 THEN 'Morning'

WHEN EXTRACT(HOUR FROM ord.order_purchase_timestamp) >= 12 AND EXTRACT(HOUR FROM ord.order_purchase_timestamp) < 17 THEN 'Afternoon'

 $WHEN\ EXTRACT(HOUR\ FROM\ ord.order_purchase_timestamp) >= 17\ AND\ EXTRACT(HOUR\ FROM\ ord.order_purchase_timestamp) < 22\ AND\ EXTRACT(HOUR\ FROM\ ord.$ THEN 'Night'

ELSE 'SLEEPING HOURS'

END AS MOMENT_DAY,

COUNT(*) as list_purchasess

FROM 'sqllproject.customers' as cust JOIN 'sqllproject.ORDERS' as ord on cust.customer_id=ord.customer_id GROUP BY MOMENT DAY ORDER BY MOMENT DAY:

```
WHEN EXTRACT(HOUR FROM ord.order_purchase_timestamp) >= 5 AND EXTRACT(HOUR FROM ord.order_purchase_timestamp) <0 THEN 'DAWN'
WHEN EXTRACT(HOUR FROM ord.order_purchase_timestamp) >= 9 AND EXTRACT(HOUR FROM ord.order_purchase_timestamp) < 12
THEN 'Morning'
WHEN EXTRACT(HOUR FROM ord.order_purchase_timestamp) >= 12 AND EXTRACT(HOUR FROM ord.order_purchase_timestamp) < 17
THEN 'Afternoon'
WHEN EXTRACT(HOUR FROM ord.order_purchase_timestamp) >= 17 AND EXTRACT(HOUR FROM ord.order_purchase_timestamp) < 22
THEN 'Night'
ELSE 'SLEENGHOURS'
SOURCE HOW ORD 'NIGHT ON THE STAND HOURS'
WHEN EXTRACT(HOUR FROM ORD.ORD 'NIGHT ON THE STAND HOURS'
SOURCE HOW ORD 'NIGHT ON THE STAND HOURS'
SOURCE HOU
sqllproject.customers' as cust
agllproject.ORDERS' as ord
customer_id=ord.customer_id
3Y MOMENT_DAY;
BY MOMENT_DAY;
```

Query results

JOB IN	IFORMATION	RESULTS		JSON	EXECUTI
Row /	MOMENT_DAY		/	list_purchasess	
1	Afternoon			32211	
2	DAWN			4888	
3	Morning			17540	
4	Night			30311	
5	SLEEPING HOUR	S		14491	

CONCLUSION: PURCHASES are maximum in the afternoon and minimum at the dawn. RECOMMENDATION:

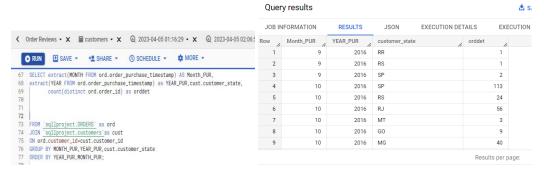
Product ads and launch should be in the afternoon more often because different users can come across it.

(3)

[a] MONTH ON MONTH orders by states

SELECT extract(MONTH FROM ord.order_purchase_timestamp) AS Month_PUR, extract(YEAR FROM ord.order_purchase_timestamp) as YEAR_PUR,cust.customer_state, count(distinct ord.order_id) as orddet

FROM `sqllproject.ORDERS` as ord JOIN `sqllproject.customers` as cust ON ord.customer_id=cust.customer_id GROUP BY MONTH_PUR,YEAR_PUR,cust.customer_state ORDER BY YEAR_PUR,MONTH_PUR;



CONCLUSION: TABLE SHOWS \mathbf{MONTH} ON \mathbf{MONTH} ORDERS by STATES.

[b] DISTRIBUTION OF CUSTOMERS ACROSS THE STATES IS:

SELECT count(customer_id),customer_state FROM `sqllproject.customers` GROUP BY customer_state

Ouary regulte

SELECT count(customer_id),customer_state FROM <u>`sqllproject.customers`</u> GROUP BY customer_state

Quei	y results				Ŭ SAVE F
JOB IN	FORMATION	RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAF
Row	f0_ //	customer_state	le		
1	485	RN			
2	1336	CE			
3	5466	RS			
4	3637	SC			
5	41746	SP			
6	11635	MG			
7	3380	BA			
8	12852	RJ			

♣ SAVE E

50

Results per page:

CONCLUSION: THESE ARE THE LIST OF CUSTOMERS SPREAD ACROSS THE STATE. SP has the highest distribution

SELECT

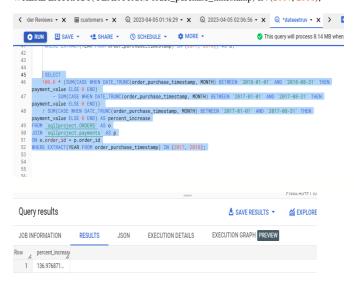
100.0 * (SUM(CASE WHEN DATE_TRUNC(order_purchase_timestamp, MONTH) BETWEEN '2018-01-01' AND '2018-08-31' THEN payment_value ELSE 0 END)
- SUM(CASE WHEN DATE_TRUNC(order_purchase_timestamp, MONTH) BETWEEN '2017-01-01' AND '2017-08-31' THEN payment_value ELSE 0 END))
/ SUM(CASE WHEN DATE_TRUNC(order_purchase_timestamp, MONTH) BETWEEN '2017-01-01' AND '2017-08-31' THEN payment_value ELSE 0 END) AS percent_increase

≛ SAVE RESULTS ▼

FROM `sqllproject.ORDERS` AS o JOIN `sqllproject.payments` AS p

ON o.order_id = p.order_id

WHERE EXTRACT(YEAR FROM order_purchase_timestamp) IN (2017, 2018);



CONCLUSION: The percentage increase is 136.9.

[b] SELECT customer_state,

Query results

TION GRAPH PREVI	DETAILS E	EXECUTION	JSON	RESULTS	FORMATION	JOB IN
	total_cost 186168.96	176.46	avg_total_cost	le	customer_state MT	Row 1
	151171.99	183.46			MA	2
	96229.4	216.73			AL	3
	5921678.12	124.8			SP	4
	1856161.49	141.38			MG	5
	322237.69	178.43			PE	6
	2129681.98	146.08			RJ	7

CONCLUSION: mean =avg(cost+freight),sum=sum(price+freight). FOR each state it will be shown.

```
customer_id,
order_status,
order_purchase_timestamp,
order_delivered_carrier_date,
order_delivered_customer_date,
order_estimated_delivery_date,
DATE_DIFF( order_delivered_customer_date,order_purchase_timestamp,DAY) AS days_to_deliver,
DATE_DIFF( order_delivered_customer_date,order_estimated_delivery_date,DAY) AS days_late
FROM `sqllproject.ORDERS
```



Que	ery results	5					≜ SAVE RESULTS ▼	盆 EXPLORE DATA	+ 0 ×
JOB I	INFORMATIO	ON RESULTS	JSON	EXECUTION DETAILS	EXECUTION GRAPH PREVIEW				
Row 308	ae4f0a	order_status delivered	6	order_purchase_timestamp 2017-12-03 18:18:26 UTC	order_delivered_carrier_date // 2017-12-07 16:39:58 UTC	order_delivered_customer_date // 2018-01-23 15:58:31 UTC	order_estimated_delivery_date 2017-12-28 00:00:00 UTC	days_to_deliver,	days_late 26
309	d0d41a	delivered		2017-12-04 12:14:59 UTC	2017-12-06 14:09:20 UTC	2018-01-04 15:28:56 UTC	2017-12-28 00:00:00 UTC	31	7
310	d0cb47	delivered		2017-11-29 13:25:18 UTC	2017-12-01 19:28:43 UTC	2017-12-28 21:05:16 UTC	2017-12-28 00:00:00 UTC	29	0
311	0a16d0	delivered		2017-11-24 16:12:54 UTC	2017-11-28 20:51:50 UTC	2018-01-03 09:39:52 UTC	2017-12-28 00:00:00 UTC	39	6
312	'4f03b9	delivered		2017-11-29 15:48:38 UTC	2017-11-30 15:27:11 UTC	2018-01-07 18:29:22 UTC	2017-12-28 00:00:00 UTC	39	10
313	/92eaca	delivered		2017-11-18 00:58:41 UTC	2017-11-27 15:42:37 UTC	2017-12-29 01:48:17 UTC	2017-12-28 00:00:00 UTC	41	1
314	896e5d	delivered		2017-12-05 15:45:05 UTC	2017-12-07 00:56:46 UTC	2018-01-04 22:49:00 UTC	2017-12-28 00:00:00 UTC	30	7
315	250dd6f	delivered		2017-11-28 12:00:43 UTC	2017-12-16 15:38:59 UTC	2018-01-03 23:35:07 UTC	2017-12-28 00:00:00 UTC	36	6

CONCLUSION:

5 (1) SELECT order_id,

days_to_deliver shows the time between purchases date and estimated delivery date. days_late shows whether the ordered was delivered within the estimated delivery date.

5[2]

```
SELECT order_id, customer_id,
DATE_DIFF(order_delivered_customer_date,order_purchase_timestamp,DAY) AS time_to_delivery,
DATE_DIFF(order_delivered_customer_date,order_estimated_delivery_date, DAY) AS diff_estimated_delivery
FROM `sqllproject.ORDERS`
WHERE order_id, customer_id,

DATE_DIFF(order_delivered_customer_date,order_purchase_timestamp ,DAY) AS time_to_delivery,
DATE_DIFF(order_delivered_customer_date,order_estimated_delivery_date, DAY) AS diff_estimated_delivery
FROM `sqllproject.ORDERS`
WHERE order_status = 'delivered';
```

Query results

JOB IN	IFORMATION	RESULTS	JSON	EXECUTION DET	AILS EXE	CUTION GRAPH	REVIEW
Row /	order_id	1	customer_id	1	time_to_delivery	diff_estimated_c	
1	635c894d068ac	37e6e03dc54e	7a34a8e89076	5ad6f90db76d0	30	-1	
2	3b97562c3aee8l	odedcb5c2e45	065d53860347	'd845788e041c	32	0	
3	68f47f50f04c4cb	6774570cfde	0378e1381c73	0d4504ebc07d2	29	-1	
4	276e9ec344d3bt	029ff83a161c	d33e520a99eb	4cfc0d3ef2b6ff	43	4	
5	54e1a3c2b97fb0	809da548a59	a0bc11375dd3	d8bdd0e0bfcbc	40	4	
6	fd04fa4105ee80	45f6a0139ca5	8fe0db7abbcc	af2d788689e91	37	1	
7	302bb8109d097	a9fc6e9cefc5	22c0028cdec9	5ad1808c1fd50	33	5	
8	66057d37308e7	87052a32828	dca924c5e55e	17bdba2ad42ae	38	6	
9	19135c945c554e	eebfd7576c73	1c7a9b908094	192a2dfae2819	36	2	
10	4493e45e7ca108	34efcd38ddeb	a1fa003a1a17	fc47164251e0e	34	0	

CONCLUSION:

IF diff_estimated_delivery is in negative then order arrived before the estimated time.

```
5[3] SELECT cust.customer_state,
    avg(DATE_DIFF(order_delivered_customer_date,order_purchase_timestamp ,DAY)) AS mean_time_to_
    avg(DATE_DIFF( order_delivered_customer_date, order_estimated_delivery_date, DAY)) AS mean_diff_estimated_delivery,
    avg(ordi.freight_value) as meanfrv
    FROM `sqllproject.customers` cust
    JOIN 'sqllproject.ORDERS' as ord
    ON cust.customer_id=ord.customer_id
    JOIN `sqllproject.Order Items` ordi
    on ord.order_id=ordi.order_id
    WHERE order_status = 'delivered'
    GROUP BY cust.customer_state
SELECT cust.customer.state,
avg(DATE_DIFF(order_delivered_customer_date,order_purchase_timestamp_,DAY)) AS mean_time_to_delivery,
avg(DATE_DIFF(order_delivered_customer_date,order_estimated_delivery_date, DAY)) AS mean_diff_estimated_delivery,
avg(ordi.freight_value) an meanfrv
avg(Date_Dispers_customer_date,out)

OR cust.customer_idend.customer_idend
OR cust.customer_idend.customer_idend
ON sallproject.Order_Idend
avg(ordi.freight_value) are customer_idend
OR cust.customer_idendi.order_idendi.order_idendi.
GROUP 8F cust.customer_state
OSOUP 8F cust.customer_state
JOSON FSGURANGION RESULTS JSON EXECUTION DETAILS EXECUTION GRAPH FRENDEN
ROW customer_state mean_time_to_st mean_tiff_estiny meanfur / 1 3798900 22 609478
     1 GO
                                              14.9481774... 11.3728590...
8.25966279... 10.2641415...
                                                                              22.5628678...
15.1151823...
                                              14.7082993...
                                                               13.2030001...
                                                                                21.6131920..
     4 BA
                                              18.7746402...
                                                              10.1194678...
                                                                                26.4875563..
                                              11.5140910...
                                                               12.3990399...
                                                                                20.6263425..
     6 MT
                                              17.5081967...
                                                               13.6393442...
                                                                                27.9969141..
     7 RJ
                                              14.6888213...
                                                               11.1396450...
                                                                                20.9114360..
Load more
```

5[5] Top 5 states with highest average freight value:

```
SELECT cust.customer_state,
```

avg(ordi.freight_value) as meanfrv FROM `sqllproject.customers` cust JOIN `sqllproject.ORDERS` as ord ON cust.customer_id=ord.customer_id JOIN `sqllproject.Order Items` ordi on ord.order_id=ordi.order_id
WHERE order_status = 'delivered' GROUP BY cust.customer_state
ORDER BY meanfrv desc
LIMIT 5

```
| avg(ordi.freight_value) as meanfrv |
| avg(ordi.freight_value) as meanfrv |
| FROM 'sqllproject.customers' cust |
| JOIN 'sqllproject.ORDERS' as ord |
| ON cust.customer_id=ord.customer_id |
| JOIN 'sqllproject.Order Items' ordi |
| on ord.order_id=ordi.order_id |
| WHERE order_status = 'delivered' |
| GROUP BY cust.customer_state |
| ORDER BY meanfrv desc |
| LIMIT 5
```

Top 5 states with lowest average freight value: SELECT cust.customer_state,

avg(ordi.freight_value) as meanfrv FROM `sqllproject.customers` cust JOIN` sqllproject.ORDERS` as ord ON cust.customer_id=ord.customer_id JOIN` sqllproject.Order Items` ordi on ord.order_id=ordi.order_id WHERE order_status = 'delivered' GROUP BY cust.customer_state ORDER BY meanfrv LIMIT 5

Query results

JOB IN	IFORMATION	RESULTS	JSON	E
Row	customer_state	le	meanfrv	
1	PB		43.0916894	
2	RR		43.0880434	
3	RO		41.3305494	
4	AC		40.0479120	
5	PI		39.1150860	

Query results

JOB IN	DB INFORMATION RESULTS		JSON	EXECUTION
Row /	customer_state	le.	meanfrv /	
1	SP		15.1151823	
2	PR		20.4718162	
3	MG		20.6263425	
4	RJ		20.9114360	
5	DF		21.0721613	

RECOMMENDATIONS: IF the freight value is to high, you might want to consider warehouses in different places to make it easy.

5[6]HIGHEST AVERAGE TIME TO DELIVERY:

SELECT cust.customer_state, avg(DATE_DIFF(order_delivered_customer_date,order_purchase_timestamp_,DAY)) AS avg_time_to_delivery,

FROM 'sqllproject.customers' cust JOIN 'sqllproject.ORDERS' as ord ON cust.customer_id=ord.customer_id JOIN `sqllproject.Order Items` ordi on ord.order_id=ordi.order_id WHERE order_status = 'delivered' GROUP BY cust.customer_state ORDER BY avg_time_to_delivery desc Limit 5



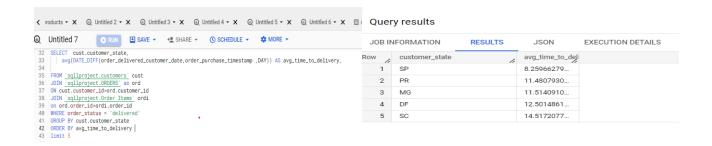
Quer	y results			
JOB IN	FORMATION	RESULTS	JSON	EXECUTION D
Row	customer_state	/	avg_time_to_del	
1	RR		27.8260869	
2	AP		27.7530864	
3	AM		25.9631901	
4	AL		23.9929742	
5	PA		23.3017077	

LOWEST AVERAGE TIME TO DELIVERY:

SELECT cust.customer_state,

 $avg(DATE_DIFF(order_delivered_customer_date, order_purchase_timestamp\ , DAY))\ AS\ avg_time_to_delivery,$

FROM 'sqllproject.customers' cust JOIN 'sqllproject.ORDERS' as ord ON cust.customer_id=ord.customer_id JOIN `sqllproject.Order Items` ordi on ord.order_id=ordi.order_id WHERE order_status = 'delivered' GROUP BY cust.customer_state
ORDER BY avg_time_to_delivery limit 5



RECOMMENDATION:

IF the delivery is slow:

Communication with customers: It is important for companies to communicate with customers about delays or delivery issues. They must provide new information about their orders and be clear about any challenges they will face.

Improve logistics: Companies can improve delivery by improving delivery processes, partnering with trusted carriers, and using tracking systems to ensure products are delivered on time.

Offer Alternative Shipping Options: If businesses experience delays while using the shipping process, they may consider offering customers alternative shipping options such as expedited shipping or same-day shipping.

Service or Return: If the customer experiences a delay or problem with delivery, the company may wish to offer a refund or refund to resolve the issue.

This helps increase customer satisfaction and loyalty.

5[7]

Top 5 states where delivery is really fast compared to estimated date:

SELECT cust.customer_state,

 $avg(DATE_DIFF(\ order_estimated_delivery_date,\ order_delivered_customer_date,DAY))\ AS\ avg_delivery_time,$

FROM `sqllproject.customers` cust JOIN `sqllproject.ORDERS` as ord ON cust.customer_id=ord.customer_id JOIN `sqllproject.Order Items` ord on ord.order_id=ordi.order_id
WHERE order_status = 'delivered' GROUP BY cust.customer_state
ORDER BY avg_delivery_time ASC limit 5:

SELECT cust.customer_state, avg(DATE_DIFF(order_estimated_delivery_date, order_delivered_customer_date, DAY)) AS avg_delivery_time, FROM _sqllproject.customers _ cust JOIN _sqllproject.ORDERS _ as ord ON cust.customer_id=ord.customer_id JOIN _sqllproject.Order_Items _ ordi on ord.order_id=ordi.order_id MHERE order_status = 'delivered' GROUP BY cust.customer_state DROER BY avg_delivery_time ASC

Query results

JOB IN	NFORMATION	RESULTS	JSON	EX
Row /	customer_state	//	avg_delivery_tim	
1	AL		7.97658079	
2	MA		9.10999999	
3	SE		9.16533333	
4	ES		9.76853932	
5	BA		10.1194678	

NOT SO FAST:

limit 5;

SELECT cust.customer_state,

avg(DATE_DIFF(order_estimated_delivery_date, order_delivered_customer_date,DAY)) AS avg_delivery_time,

FROM `sqllproject.customers` cust JOIN `sqllproject.ORDERS` as ord ON cust.customer_id=ord.customer_id JOIN `sqllproject.Order Items` ordi on ord.order_id=ordi.order_id WHERE order_status = 'delivered' GROUP BY cust.customer_state ORDER BY avg_delivery_time DESC limit 5;

Query results

JOB IN	NFORMATION RESULTS		JSON	EXEC
Row /	customer_state	6	avg_delivery_tim	
1	AC		20.0109890	
2	RO		19.0805860	
3	AM		18.9754601	
4	AP		17.4444444	
5	RR		17.4347826	

```
[6]
[a] MONTH OVER MONTH:
SELECT *,
order_count-LAG(order_count,1,0) OVER (PARTITION BY payment_type, year Order By month) as monthomonth
FROM

(SELECT EXTRACT(YEAR from order_purchase_timestamp) as YEAR,
EXTRACT(MONTH FROM order_purchase_timestamp) AS month,
p.payment_type,
COUNT(DISTINCT o.order_id) AS order_count
FROM

'sqllproject.ORDERS' o
JOIN 'sqllproject.payments' p ON o.order_id = p.order_id
GROUP BY
1,2,3) as lol
ORDER BY
1,2,3.
SELECT *,
order_count-LAG(order_count,1,0) OVER (PARTITION BY payment_type, year Order By month) as monthomonth
FROM

EXTRACT(WEAR from order_purchase_timestamp) as YEAR,
EXTRACT(WEAR from order_purchase_timestamp) AS month,
p.payment_type,
COUNT(DISTINCT o.order_id) AS order_count
FROM

'sqllproject.ORDERS' o
JOIN 'sqllproject.payments' p ON o.order_id = p.order_id
GROUP BY
1,2,3) as lol
GROUP BY
1,2,3 as lol
```

Query results

JOB IN	IFORMATION	RESULTS	JSON	EXECUTION DET	TAILS EXE	CUTION GRAPH PREVIO
Row /	YEAR //	month //	payment_type	h	order_count //	monthomonth /
3	2016	10	credit_card		253	250
4	2016	10	debit_card		2	2
5	2016	10	voucher		11	11
6	2016	12	credit_card		1	-252
7	2017	1	UPI		197	197
8	2017	1	credit_card		582	582
9	2017	1	debit_card		9	9
10	2017	1	voucher		33	33
11	2017	2	UPI		398	201
12	2017	2	credit_card		1347	765

CONCLUSION:

THE month over month is shown by monthomonth with respect to different years.

```
6[b] COUNT of ORDERS based on no. of payment installments:
SELECT
 pay.payment_installments,
 COUNT(DISTINCT ordi.order_id) AS OrderCount
FROM
  sqllproject.ORDERS` ordi
 JOIN `sqllproject.payments` pay ON ordi.order_id = pay.order_id
GROUP BY
 pay.payment_installments
ORDER BY
 pay.payment_installments 5
                      ☑ SAVE ▼ + SHARE ▼ ⑤ SCHEDULE ▼
  SELECT
  pay.payment_installments,
COUNT(DISTINCT ordi.order_id) AS OrderCount
FROM
  sqllproject.ORDERS` ordi
JOIN 'sqllproject.payments` pay ON ordi.order_id = pay.order_id
GROUP BY
  pay.payment_installments
ORDER BY
    pay.payment_installments
```

Query results

JOB INFORMATION		RESULTS	JSON	
Row ./	payment_installr	OrderCount _//		
2	1.	49060		
3	2	12389		
4	3	10443		
5	4	7088		
6	5	5234		
7	6	3916		
8	7	1623		
9	8	4253		

CONCLUSION: BASED on each payment installments the count of orders has been shown above.

(7) The actionable insights are shown in the conclusion of each questions.

(8)RECOMMENDATION GIVEN AFTER FEW QUESTIONS. However here are my overall analysis: Based on the analysis of the e-commerce dataset for Brazil, there are several recommendations that can be made to improve the company's operations:

Improve the logistics. The analysis reveals that there are some issues with delivery times in certain regions, so the company could consider optimizing its shipping processes or partnering with reliable shipping carriers to improve its delivery times.

Focus on marketing during peak months. The analysis suggests that there are some seasonality trends in e-commerce sales in Brazil, with peaks during specific months. The company could focus on marketing during these peak months to capitalize on the increased demand.

Improve estimated delivery times. The analysis reveals that there are some issues with estimated delivery times not being met, which could lead to dissatisfied customers. By improving the accuracy of its estimated delivery times, the company could improve customer satisfaction and loyalty.