SQL PROJECT

select \* from olist\_customers\_dataset;

select \* from olist\_order\_items\_dataset;

select \* from olist\_order\_payments\_dataset;

select \* from olist\_orders\_dataset;

select \* from olist\_products\_dataset;

select \* from olist\_sellers\_dataset;

select \* from product\_category\_name\_translation;

select\* from olist\_order\_reviews\_dataset

/\*a. Create the different metrics like Sales, customer acquisitions, total no. of orders

for each Year across the different states they serve.

Does all the metrics show similar trends or is there any disparity amongst each of them?

\*/

select distinct seller\_state,year,sum(price) as price from

(select DATEPART(year,shipping\_limit\_date) as year,seller\_state, price

from olist\_order\_items\_dataset as a left join olist\_sellers\_dataset as b on

a.seller\_id=b.seller\_id

group by DATEPART(year,shipping\_limit\_date),seller\_state,price) as f

group by seller\_state,year

order by year;

select DATEPART(year,order\_approved\_at) as year,customer\_state,count(a.customer\_id) as total\_ID

from olist\_customers\_dataset as a join olist\_orders\_dataset as b

on a.customer\_id=b.customer\_id

group by DATEPART(year,order\_approved\_at),customer\_state

order by year;

select DATEPART(year,order\_approved\_at) as year,customer\_state,count(order\_id) as total\_order\_ID

from olist\_customers\_dataset as a join olist\_orders\_dataset as b

on a.customer\_id=b.customer\_id

group by DATEPART(year,order\_approved\_at),customer\_state

order by year;

/\*b. Using the above metrics, identify the top 2 States which show

i. Declining trend over the years

ii. Increasing trend over the years

\*/

Increasing trend

select distinct seller\_state from

(select seller\_state,year,price,rnk from

(select \*,dense\_rank() over(partition by seller\_state order by price asc) as rnk from

(select seller\_state , year,sum(price) as price from

(select DATEPART(year,shipping\_limit\_date) as year,seller\_state, price

from olist\_order\_items\_dataset as a left join olist\_sellers\_dataset as b on

a.seller\_id=b.seller\_id

group by DATEPART(year,shipping\_limit\_date),seller\_state,price) as f

group by seller\_state,YEAR) g)k

where rnk<=3 and price>190000

) n

order by seller\_state desc;

Decreasing trend

select seller\_state from

(select \* from

(select \*,dense\_rank() over(partition by seller\_state order by price desc) as rnk from

(select seller\_state , year,sum(price) as price from

(select DATEPART(year,shipping\_limit\_date) as year,seller\_state, price

from olist\_order\_items\_dataset as a left join olist\_sellers\_dataset as b on

a.seller\_id=b.seller\_id

group by DATEPART(year,shipping\_limit\_date),seller\_state,price) as f

group by seller\_state,YEAR) g)k

where rnk<=3 and price<200) s

where price>160

order by seller\_state desc ;

Count customerid for PR,SP

SELECT customer\_state, COUNT(CUSTomer\_id) as total\_customerID from

(select customer\_id,customer\_city,customer\_state from olist\_customers\_dataset

where customer\_state IN ('SP','PR')) AS F

group by customer\_state

Count customerid for PR,SP for city

SELECT customer\_state,Customer\_city, COUNT(CUSTomer\_id) as total\_customerID from

(select customer\_id,customer\_city,customer\_state from olist\_customers\_dataset

where customer\_state IN ('SP','PR')) AS F

group by customer\_state,customer\_city

Count OrderID

select customer\_state,count(order\_id) as total\_orderid

from (select a.customer\_id,order\_id,customer\_city,customer\_state

from olist\_customers\_dataset as a join

olist\_orders\_dataset as b on a.customer\_id=

b.customer\_id) as c

group by customer\_state

order by count(order\_id) desc;

Type of payment mode and there total

Select distinct customer\_state,

sum(case when payment\_type= 'credit\_card' then payment\_installments end) as 'credit\_card',

sum(case when payment\_type= 'boleto' then payment\_installments end) as 'boleto',

sum(case when payment\_type= 'voucher' then payment\_installments end) 'voucher',

sum(case when payment\_type= 'debit\_card' then payment\_installments end) 'debit\_card',

sum(case when payment\_type= 'not\_defined' then payment\_installments end) 'not\_defined'

from olist\_customers\_dataset as a join olist\_orders\_dataset as c on

a.customer\_id=c.customer\_id

join olist\_order\_payments\_dataset as b

on b.order\_id=c.order\_id

where customer\_state IN ('PR','SP','RO','BA')

group by customer\_state;

Average Order Date

select customer\_state, avg(DATEDIFF(day,order\_delivered\_carrier\_date,order\_delivered\_customer\_date)) as days from

olist\_customers\_dataset as a join olist\_orders\_dataset as b on a.customer\_id=b.customer\_id

where customer\_state in ('SP','PR','RO','BA')

group by customer\_state;