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 metrices show similar trends or is there any disparity amongst each of them?

```
--A) Sales
WITH CTE1 AS(
SELECT YEAR(0.order_purchase_timestamp)Years,C.customer_state,ROUND(SUM(I.price),2)Sales
FROM customers C
JOIN order1 0
ON C.customer id = O.customer id
JOIN order items I
ON 0.order_id = I.order_id
WHERE O.order_status NOT IN('unavailable','canceled')
GROUP BY C.customer_state, YEAR(0.order_purchase_timestamp)),
CTE2 AS(
SELECT *, ROW_NUMBER() OVER(PARTITION BY Years, customer_state ORDER BY
customer_state)Ranks
FROM CTE1)
SELECT Years, customer state, sales FROM CTE2;
   SQL Project 1.sql - DESKTOP-Q54KQ0R.Project1 (DESKTOP-Q54KQ0R\welcome (58)) - Microsoft SQL Ser
                                                                                                                          Quick Launch (Ctrl+Q)
                                                                                                                                             P _ 8 X
      Edit View Query Project Tools Window
    ○ - ○ 15 - 13 - 4 単 3 A New Query 通 最 最 最 最 最 表 日 白 ? - で - 図 - 声
                                                                                    · 5 / = 5 · -
    ₩ ₩ | Project1
                       · ▶ Execute = ✔ 80 日日 80 88 日 日田日 3 3 조조 10 10 1
    Connect * # *# = 7 C **
                               RITH CTE1 AS(
SELECT YEAR(O.order_purchase_timestamp)Years,C.customer_state_#00HO(SUM(I.price),2)Sales
FROM Customers C
100H order1 0
0H C.customer_id = 0.customer_id
    ■ DESKTOP-Q54KQ0R (SQL Server 15.0.20
     Replication
      PolyBase

    Olybase
    Always On High Availability
    Management
    Integration Services Catalogs
    SQL Server Agent (Agent XPs disable)

                                OIN order_items I
ON O.order_id = I.order_id
                                 ERE O.order_status NOT IN('unavailable','canceled')
                                  DUP BY C.customer_state, YEAR(O.order_purchase_timestamp)),
                                       DW_NUMBER() OVER(PARTITION BY Years customer_state ORDER BY customer_state)Ranks
                                FROM CTE1)
                             ⊞ Results 🖼 Messages
                                  🚈 O 터 🧑 🕿 🕖 🐠 🔘 🕫 🚍 📵 🧶 🗷
   Type here to search
--B) Customer Acquisitions
```

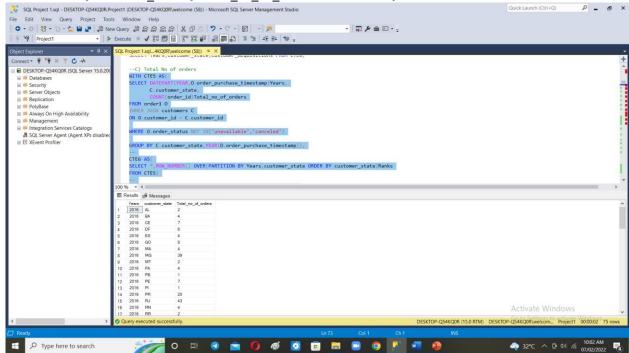
WITH CTE3 AS(

```
SELECT
YEAR(O.order purchase timestamp)Years, C. customer state, COUNT(C. customer id)Customer Acqui
sitions FROM customers C
INNER JOIN order1 0
ON C.customer id = O.customer id
WHERE O.order status NOT IN('unavailable','canceled')
GROUP BY C.customer_state, YEAR(O.order_purchase_timestamp)),
CTE4 AS(
SELECT *,ROW_NUMBER() OVER(PARTITION BY Years,customer_state ORDER BY
customer state)Ranks
FROM CTE3)
SELECT Years,customer_state,Customer_Acquisitions FROM CTE4;
                                                                                                  Quick Launch (Ctrl+Q)
                                                                                                                 P _ 5 ×
  SQL Project 1.sql - DESKTOP-Q54KQ0R.Project1 (DESKTOP-Q54KQ0R\v
   ○ - ○ | 15 - 15 - 4 | 12 | 12 | 12 | New Query 通 点 点 点 点 点 17 - 17 - 18 | - 1 | 19
                                                                    · 5 / = 0 · .
                   QL Project 1.sql...4KQ0R\welco
   WITH CTE3 AS( | SELECT YEAR(O.order_purchase_timestamp)Years,C.customer_state;COUNT(C.customer_id)Customer_Acquisitions
                          ROM customers C
INNER JOIN order1 0
N C.customer_id = O.customer_id
                           HERE O.order_status NOT IN('unavailable','canceled')
                           OUP BY C.customer_state, YEAR(O.order_purchase_timestamp)),
                                (_NUMBER() OVER(PARTITION BY Years,customer_state ORDER BY customer_state)Ranks
                           OM CTE3)
                          SELECT Years, customer_state, Customer_Acquisitions FROM CTE4;
                          --C) Total No of orders
                       Results gill Messages
                                                                                                    Activate Windows
                          O H 3 🕿 🚺 Ø 🔯 🗓 🟗 🖺 🧿 🗗 🚜 🔞
  Type here to search
--C) Total No of orders
WITH CTE5 AS(
SELECT DATEPART(YEAR, 0. order_purchase_timestamp)Years,
         C.customer state,
             COUNT(order_id)Total_no_of_orders
FROM order1 0
INNER JOIN customers C
ON O.customer_id = C.customer_id
WHERE O.order_status NOT IN('unavailable','canceled')
GROUP BY C.customer state, YEAR(O.order purchase timestamp)),
CTE6 AS(
```

```
SELECT *,ROW_NUMBER() OVER(PARTITION BY Years,customer_state ORDER BY
customer_state)Ranks
FROM CTE5)
```

--

SELECT Years,customer\_state,Total\_no\_of\_orders FROM CTE6;



C. For the States identified above, do the Root Cause analysis for their performance across a variety of metrics.

You can utilize the following metrics and explore a few yourself as well by analyzing the data.

# i)Category level Sales and orders placed,

A) INCREASING

```
SELECT
YEAR(O.order_purchase_timestamp)Years,C.customer_state,P.product_category_name,COUNT(P.pr
oduct_category_name)Order_palced
FROM order1 O
INNER JOIN customers C
ON O.customer_id = C.customer_id

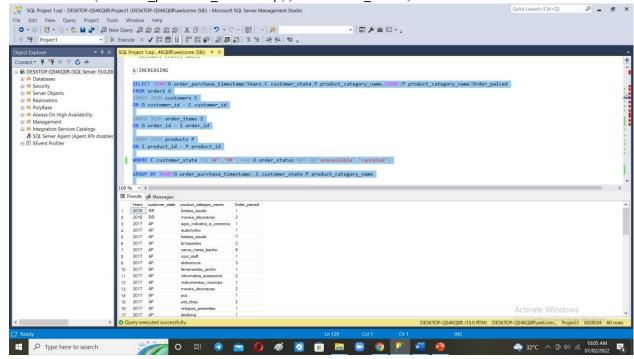
INNER JOIN order_items I
ON O.order_id = I.order_id

INNER JOIN products P
ON I.product_id = P.product_id

WHERE C.customer_state IN('AP','RR') And O.order_status NOT IN('unavailable','canceled')
```

GROUP BY YEAR(0.order\_purchase\_timestamp),C.customer\_state,P.product\_category\_name

ORDER BY YEAR(O.order purchase timestamp), C.customer state;



#### B) DECREASING

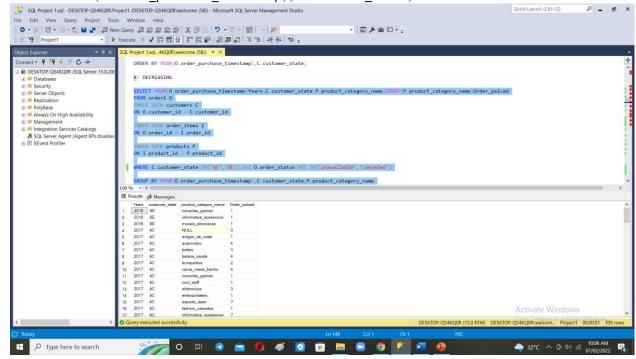
```
YEAR(O.order_purchase_timestamp)Years,C.customer_state,P.product_category_name,COUNT(P.pr
oduct_category_name)Order_palced
FROM order1 0
INNER JOIN customers C
ON O.customer_id = C.customer_id

INNER JOIN order_items I
ON O.order_id = I.order_id

INNER JOIN products P
ON I.product_id = P.product_id

WHERE C.customer_state IN('AC','SE') And O.order_status NOT IN('unavailable','canceled')
GROUP BY YEAR(O.order_purchase_timestamp),C.customer_state,P.product_category_name
```

ORDER BY YEAR(O.order purchase timestamp), C. customer state;



# 2)post-order reviews

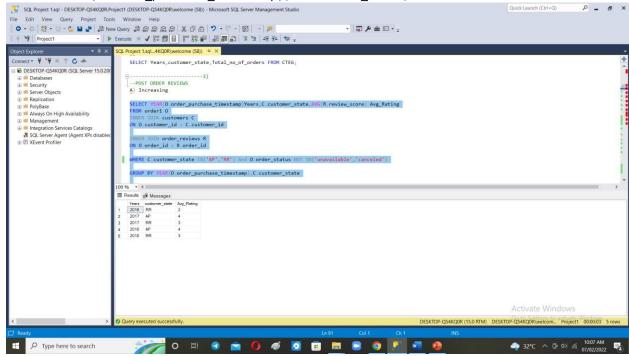
# A) Increasing

```
SELECT YEAR(0.order_purchase_timestamp)Years,C.customer_state,AVG(R.review_score)
Avg_Rating
FROM order1 0
INNER JOIN customers C
ON 0.customer_id = C.customer_id

INNER JOIN order_reviews R
ON 0.order_id = R.order_id

WHERE C.customer_state IN('AP','RR') And 0.order_status NOT IN('unavailable','canceled')
GROUP BY YEAR(0.order_purchase_timestamp),C.customer_state
```

ORDER BY YEAR(0.order\_purchase\_timestamp),C.customer\_state;



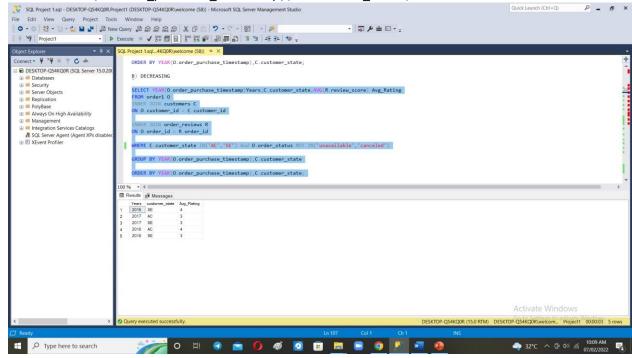
#### B) DECREASING

```
SELECT YEAR(0.order_purchase_timestamp)Years,C.customer_state,AVG(R.review_score)
Avg_Rating
FROM order1 0
INNER JOIN customers C
ON 0.customer_id = C.customer_id

INNER JOIN order_reviews R
ON 0.order_id = R.order_id

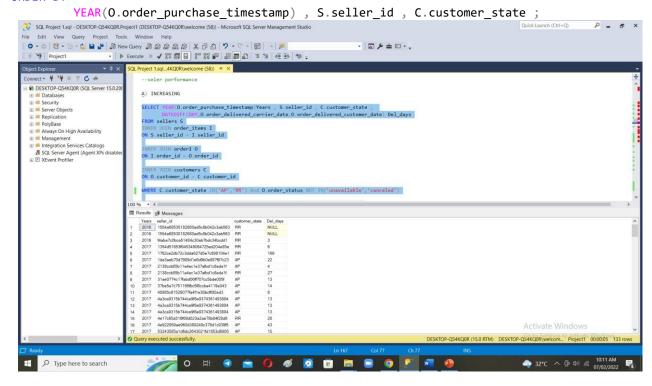
WHERE C.customer_state IN('AC','SE') And 0.order_status NOT IN('unavailable','canceled')
GROUP BY YEAR(0.order_purchase_timestamp),C.customer_state
```

ORDER BY YEAR(0.order purchase timestamp),C.customer state;



# 3)Seller performance in terms of deliveries,

# A) INCREASING



# B) DECREASING

```
ORDER BY YEAR(O.order purchase timestamp), C.customer state;
WHERE C.customer_state IN('AC','SE') And O.order_status NOT IN('unavailable','canceled')
YEAR(O.order_purchase_timestamp) , S.seller_id , C.customer_state ;
  SQL Project 1.sql - DESKTOP-Q54KQ0R.Project1 (DESKTOP-Q54KQ0R\welcome (58)) - Microsoft SQL Server Management Studio
                                                                                                                                                    Quick Launch (Ctrl+Q)
                                                                                                                                                                         P _ 6 ×
  ○ - ○ | 数 - 1 - 4 単 4 A New Query 過 品 品 品 品 は 日 白 | フ - ୧ - | 図 | - | #
                                                                                                     · | 57 / = 10 · .
                     Project1
   Connect ▼ ¥ ■ ▼ C →
                                     ORDER BY YEAR(O.order_purchase_timestamp) , S.seller_id , C.customer_state ;

    ■ DESKTOP-Q54KQ0R (SQL Server 15.0.2)

    Security
    Server Objects
    Replication
    PolyBase
    Always On High Availability
    Management
                                     B) DECREASING
                                     SELECT YEAR(O.order_purchase_timestamp)Years , S.seller_id , C.customer_state ,
DATEDIFF(DAY,O.order_delivered_carrier_date,O.order_delivered_customer_date) Del_days
                                      FROM sellers S
                                      INNER JOIN order_items I
ON S.seller_id = I.seller_id

    ★ Management
    ★ Integration Services Catalogs
    ★ SQL Server Agent (Agent XPs disable)
    ▼ XEvent Profiler
                                      MHERE C.customer_state IN('AC','SE') And O.order_status NOT IN('unavailable','canceled')
                                    Years seller_id 
2016 b335c59ab742f751a85db9c411a86739 
2016 ce27a3cc3c8cc1ea79d11e561e9bebb6
                                 O H @ 🕿 🕖 Ø 🔘 🖫 🚍 🧿 🥬 🗸
  Type here to search
```

# 4)product-level sales & orders placed,

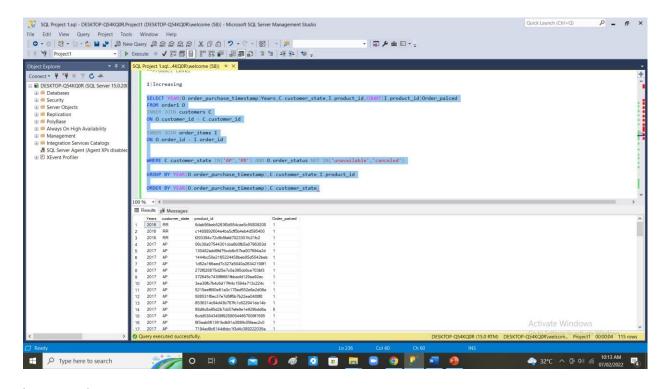
#### 1) Increasing

```
SELECT
```

```
YEAR(0.order_purchase_timestamp)Years,C.customer_state,I.product_id,COUNT(I.product_id)Or
der_palced FROM order1 0
INNER JOIN customers C
ON 0.customer_id = C.customer_id

INNER JOIN order_items I
ON 0.order_id = I.order_id

WHERE C.customer_state IN('AP','RR') AND 0.order_status NOT IN('unavailable','canceled')
GROUP BY YEAR(0.order_purchase_timestamp),C.customer_state,I.product_id
```



# 2) Decreasing

```
SELECT
YEAR(0.order_purchase_timestamp)Years,C.customer_state,I.product_id,COUNT(I.product_id)Or
der_palced FROM order1 0
INNER JOIN customers C
ON 0.customer_id = C.customer_id

INNER JOIN order_items I
ON 0.order_id = I.order_id

WHERE C.customer_state IN('AC','SE') AND 0.order_status NOT IN('unavailable','canceled')
GROUP BY YEAR(0.order_purchase_timestamp),C.customer_state,I.product_id
```

ORDER BY YEAR(O.order purchase timestamp), C.customer state; Quick Launch (Ctrl+Q) SQL Project 1.sql - DESKTOP-Q54KQ0R.Project1 (DESKTOP-Q54KQ0R\welcome (58)) - Microsoft SQL Server Man P \_ 5 × · 5 / = 0 · -Connect • • • • • • • • • DESKTOP-Q54KQ0R (SQL Server 15.0.2 SELECT YEAM(O.order\_purchase\_timestamp)Years,C.customer\_state,I.product\_id,COUNT(I.product\_id)Order\_palced
FROM order1 0
IMMER JOIN customers C
ON O.customer\_id = C.customer\_id ⊕ ■ Security

⊕ ■ Server Objects

⊕ ■ Replication PolyBase Always On High Availability Management HERE C.customer\_state IN('AC', 'SE') AND O.order\_status NOT IN('unavailable', 'canceled') ROUP BY YEAR (O order purchase timestamp) C customer state I product id ORDER BY YEAR(O.order\_purchase\_timestamp),C.customer\_state; 40e8b425d1a26e2d9cb77363523e05ce o h a 💿 🐧 🦸 🗓 🟗 🗎 🕥 🚺 📶 Type here to search 32°C ∧ © Φ) € 07/02/2022

# 5)% of orders delivered earlier than the expected date,

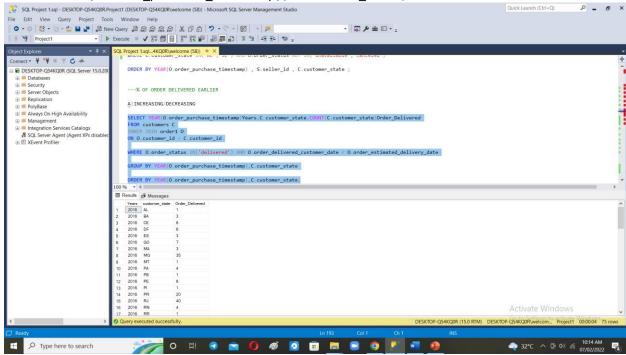
A) INCREASING/DECREASING

```
SELECT
YEAR(O.order_purchase_timestamp)Years,C.customer_state,COUNT(C.customer_state)Order_Deliv
ered
FROM customers C
INNER JOIN order1 0
ON O.customer_id = C.customer_id

WHERE O.order_status IN('delivered') AND O.order_delivered_customer_date <
O.order_estimated_delivery_date

GROUP BY YEAR(O.order_purchase_timestamp),C.customer_state</pre>
```

ORDER BY YEAR(0.order\_purchase\_timestamp), C.customer\_state;



# 6)% of orders delivered later than the expected date, etc.e

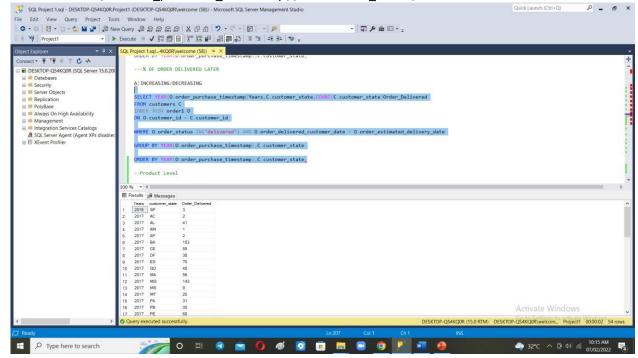
A) INCREASING/DECREASING

```
SELECT
YEAR(0.order_purchase_timestamp)Years,C.customer_state,COUNT(C.customer_state)Order_Deliv
ered
FROM customers C
INNER JOIN order1 0
ON 0.customer_id = C.customer_id

WHERE 0.order_status IN('delivered') AND 0.order_delivered_customer_date >
0.order_estimated_delivery_date

GROUP BY YEAR(0.order_purchase_timestamp),C.customer_state
```

ORDER BY YEAR(0.order purchase timestamp),C.customer state;

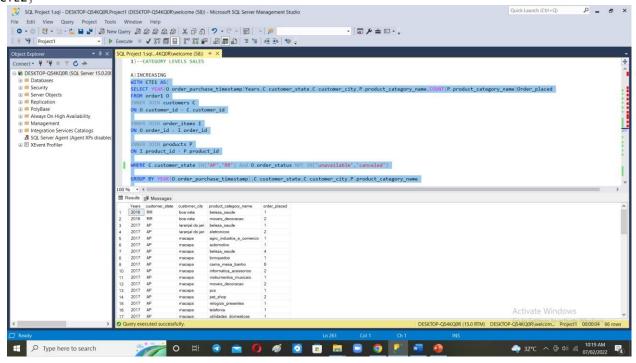


D. Do the above analysis for the top 2 cities which are causing the trend for each of the states identified in point (b)

```
1)Category level Sales and orders placed,
A) INCREASING
WITH CTE1 AS( SELECT
YEAR(O.order_purchase_timestamp)Years,C.customer_state,C.customer_city,P.product_category
_name, COUNT(P.product_category_name)Order_placed
FROM order1 0
INNER JOIN customers C
ON O.customer_id = C.customer_id
INNER JOIN order items I
ON 0.order_id = I.order_id
INNER JOIN products P
ON I.product_id = P.product_id
WHERE C.customer_state IN('AP','RR') And O.order_status NOT IN('unavailable','canceled')
GROUP BY
YEAR(O.order_purchase_timestamp), C.customer_state, C.customer_city, P.product_category_name
),
CTE2 AS(
SELECT * , DENSE_RANK() OVER(PARTITION BY Years, customer_state ORDER BY Order_placed
DESC) Ranks
```

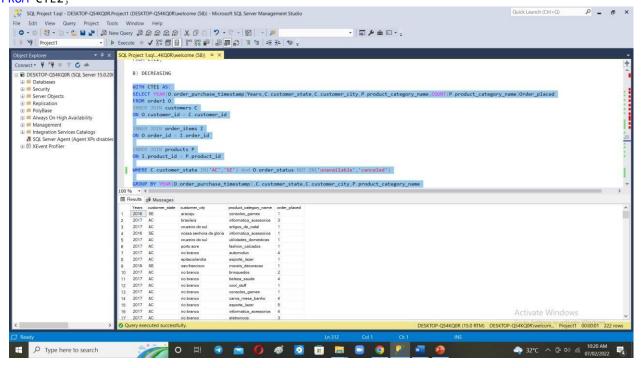
ORDER BY YEAR(0.order\_purchase\_timestamp),C.customer\_state;
FROM CTE1)

SELECT Years,customer\_state,customer\_city,product\_category\_name,order\_placed FROM
CTE2:



# B) DECREASING WITH CTE1 AS( SELECT YEAR(O.order\_purchase\_timestamp)Years,C.customer\_state,C.customer\_city,P.product\_category \_name,COUNT(P.product\_category\_name)Order\_placed FROM order1 0 INNER JOIN customers C ON O.customer\_id = C.customer\_id INNER JOIN order items I ON O.order id = I.order id INNER JOIN products P ON I.product\_id = P.product\_id WHERE C.customer\_state IN('AC', 'SE') And O.order\_status NOT IN('unavailable', 'canceled') **GROUP BY** YEAR(O.order\_purchase\_timestamp), C.customer\_state, C.customer\_city, P.product\_category\_name ), CTE2 AS( SELECT \* , DENSE RANK() OVER(PARTITION BY Years, customer state ORDER BY Order placed **DESC**) Ranks FROM CTE1)

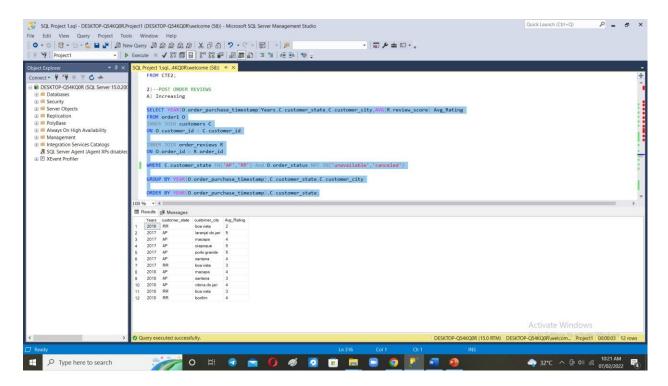
SELECT Years, customer\_state, customer\_city, product\_category\_name, order\_placed FROM CTE2;



# 2) post-order reviews,

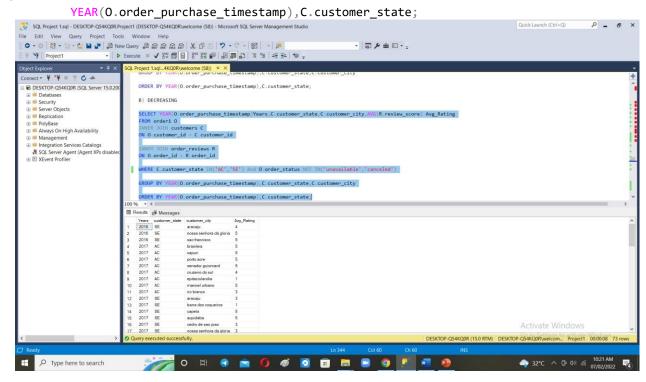
A) Increasing

```
YEAR(O.order purchase timestamp)Years, C. customer state, C. customer city, AVG(R. review score
) Avg_Rating
FROM order1 0
INNER JOIN customers C
ON O.customer id = C.customer id
INNER JOIN order reviews R
ON 0.order_id = R.order_id
WHERE C.customer_state IN('AP','RR') And O.order_status NOT IN('unavailable','canceled')
GROUP BY YEAR(0.order_purchase_timestamp), C. customer_state, C. customer_city
YEAR(0.order_purchase_timestamp),C.customer_state;
```



## B) DECREASING

```
YEAR(O.order_purchase_timestamp)Years,C.customer_state,C.customer_city,AVG(R.review_score
) Avg_Rating
FROM order1 0
INNER JOIN customers C
ON O.customer_id = C.customer_id
INNER JOIN order_reviews R
ON O.order_id = R.order_id
WHERE C.customer_state IN('AC', 'SE') And O.order_status NOT IN('unavailable', 'canceled')
GROUP BY YEAR(O.order purchase timestamp), C.customer state, C.customer city
```

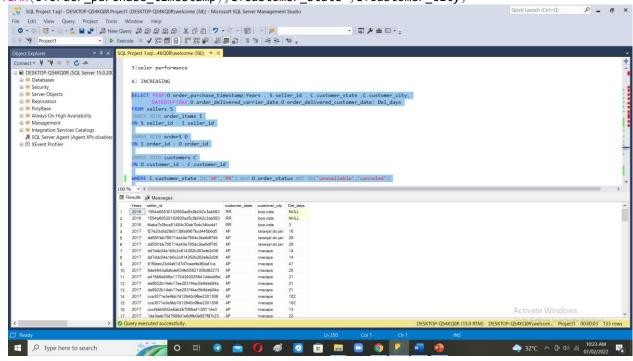


# 3)Seller performance in terms of deliveries,

A) INCREASING

ON O.customer\_id = C.customer\_id

WHERE C.customer\_state IN('AP','RR') And O.order\_status NOT IN('unavailable','canceled')
YEAR(O.order\_purchase\_timestamp),C.customer\_state ,C.customer\_city;



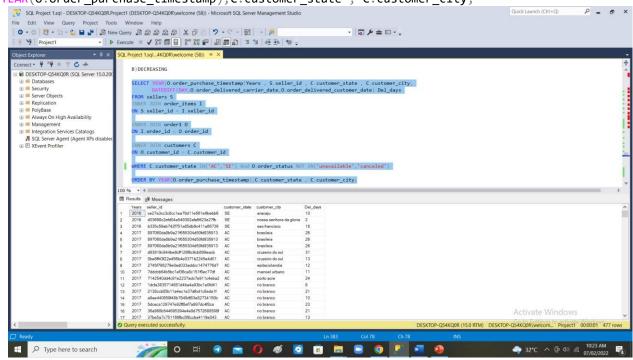
## B) DECREASING

ON I.order\_id = 0.order\_id

INNER JOIN customers C

ON O.customer id = C.customer id

WHERE C.customer\_state IN('AC','SE') And O.order\_status NOT IN('unavailable','canceled')
YEAR(O.order\_purchase\_timestamp), C.customer\_state , C.customer\_city;



# 4)product-level sales & orders placed,

1)Increasing

#### **SELECT**

YEAR(O.order\_purchase\_timestamp)Years,C.customer\_state,C.customer\_city,I.product\_id,COUNT
(I.product\_id)Order\_palced

FROM order1 0

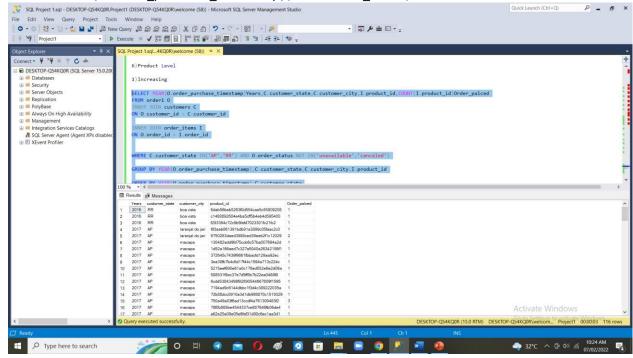
INNER JOIN customers C

```
ORDER BY
ON O.customer_id = C.customer_id

INNER JOIN order_items I
ON O.order_id = I.order_id

WHERE C.customer_state IN('AP','RR') AND O.order_status NOT IN('unavailable','canceled')
GROUP BY YEAR(O.order_purchase_timestamp),C.customer_state,C.customer_city,I.product_id
```

ORDER BY YEAR(O.order purchase timestamp), C. customer state;



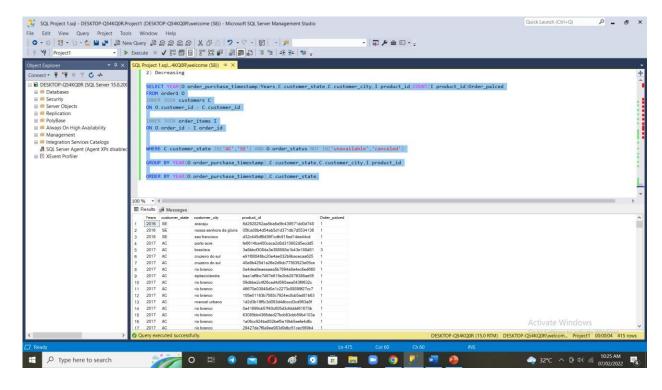
#### 2) Decreasing

```
YEAR(0.order_purchase_timestamp)Years,C.customer_state,C.customer_city,I.product_id,COUNT
(I.product_id)Order_palced
FROM order1 0
INNER JOIN customers C
ON 0.customer_id = C.customer_id

INNER JOIN order_items I
ON 0.order_id = I.order_id

WHERE C.customer_state IN('AC','SE') AND 0.order_status NOT IN('unavailable','canceled')
GROUP BY YEAR(0.order_purchase_timestamp),C.customer_state,C.customer_city,I.product_id
```

ORDER BY YEAR(O.order\_purchase\_timestamp), C.customer\_state;



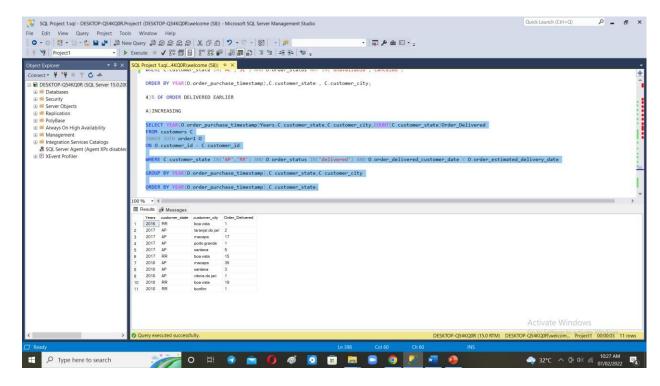
# 5)% of orders delivered earlier than the expected date, ${\tt A)\,INCREASING}$

```
YEAR(0.order_purchase_timestamp)Years,C.customer_state,C.customer_city,COUNT(C.customer_s
tate)Order_Delivered FROM customers C
INNER JOIN order1 0
ON 0.customer_id = C.customer_id

WHERE C.customer_state IN('AP','RR') AND 0.order_status IN('delivered') AND
0.order_delivered_customer_date < 0.order_estimated_delivery_date

GROUP BY YEAR(0.order_purchase_timestamp),C.customer_state,C.customer_city</pre>
```

# ORDER BY YEAR(0.order\_purchase\_timestamp),C.customer\_state;

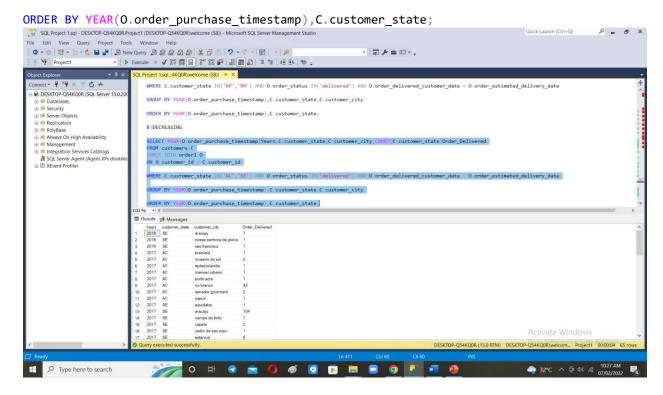


# B) DECREASING

```
YEAR(0.order_purchase_timestamp)Years,C.customer_state,C.customer_city,COUNT(C.customer_s
tate)Order_Delivered FROM customers C
INNER JOIN order1 0
ON O.customer_id = C.customer_id

WHERE C.customer_state IN('AC','SE') AND O.order_status IN('delivered') AND
O.order_delivered_customer_date < O.order_estimated_delivery_date

GROUP BY YEAR(O.order_purchase_timestamp),C.customer_state,C.customer_city
```



## 6)% of orders delivered later than the expected date, etc.e

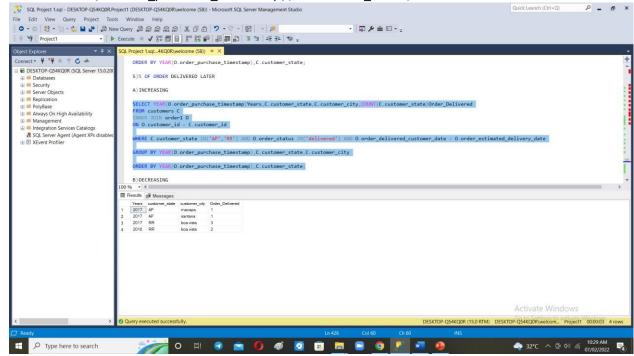
# A) INCREASING

```
YEAR(0.order_purchase_timestamp)Years,C.customer_state,C.customer_city,COUNT(C.customer_s
tate)Order_Delivered FROM customers C
INNER JOIN order1 0
ON 0.customer_id = C.customer_id

WHERE C.customer_state IN('AP','RR') AND 0.order_status IN('delivered') AND
0.order_delivered_customer_date > 0.order_estimated_delivery_date

GROUP BY YEAR(0.order_purchase_timestamp),C.customer_state,C.customer_city
```

ORDER BY YEAR(0.order\_purchase\_timestamp),C.customer\_state;



#### B) DECREASING

```
YEAR(0.order_purchase_timestamp)Years,C.customer_state,C.customer_city,COUNT(C.customer_s
tate)Order_Delivered FROM customers C
INNER JOIN order1 0
ON 0.customer_id = C.customer_id

WHERE C.customer_state IN('AC','SE') AND 0.order_status IN('delivered') AND
0.order_delivered_customer_date > 0.order_estimated_delivery_date

GROUP BY YEAR(0.order_purchase_timestamp),C.customer_state,C.customer_city
```

ORDER BY YEAR(O.order purchase timestamp), C. customer state;

