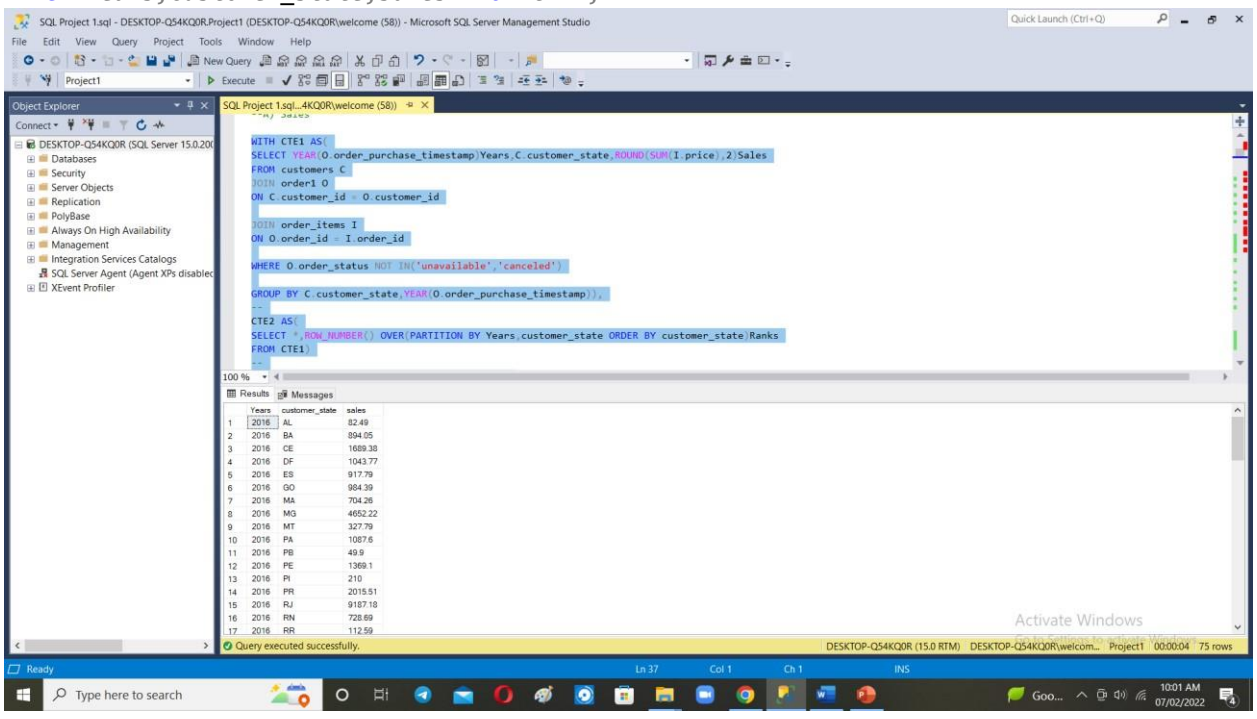
A. Create the different metrics like Sales, customer acquisitions, total no. of orders for each Year across the different states they serve.

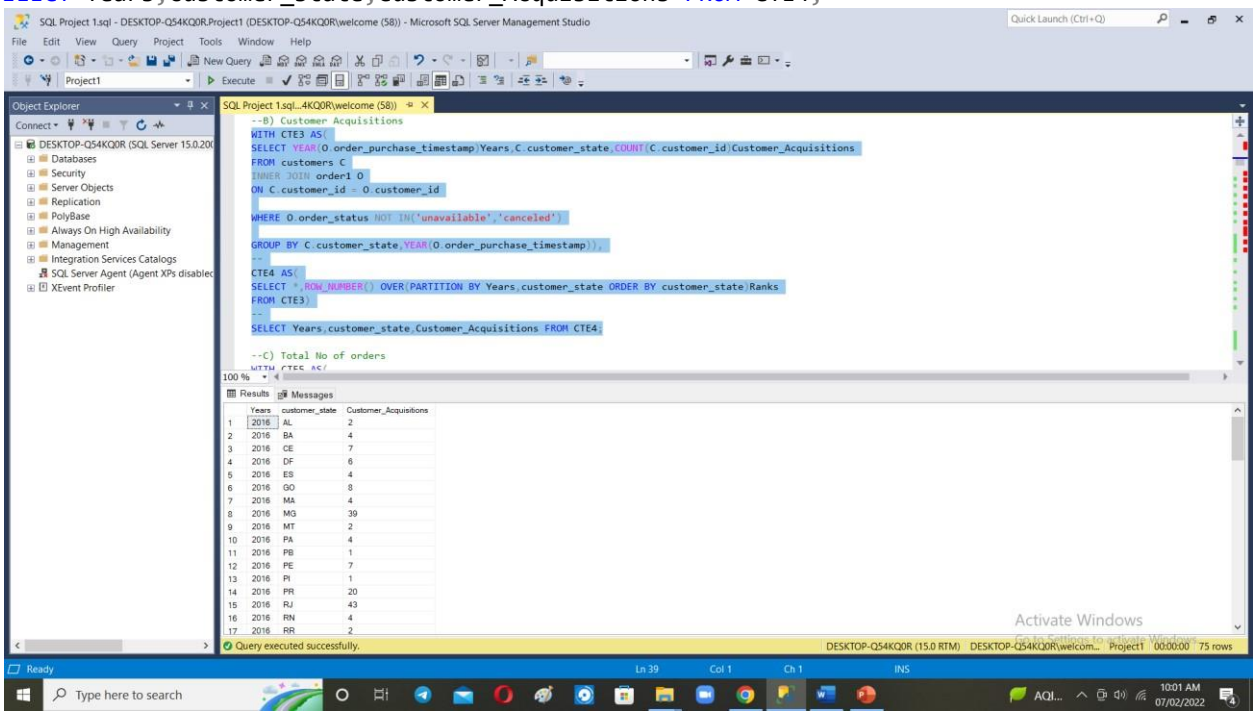
Does all the metrices show similar trends or is there any disparity amongst each of them?

```
--A) Sales

WITH CTE1 AS(
SELECT YEAR(O.order_purchase_timestamp)Years,C.customer_state,ROUND(SUM(I.price),2)Sales
FROM customers C
JOIN order1 O
ON C.customer_id = O.customer_id

JOIN order_items I
ON O.order_id = I.order_id

WHERE O.order_status NOT IN('unavailable','canceled')

GROUP BY C.customer_state,YEAR(O.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;
```



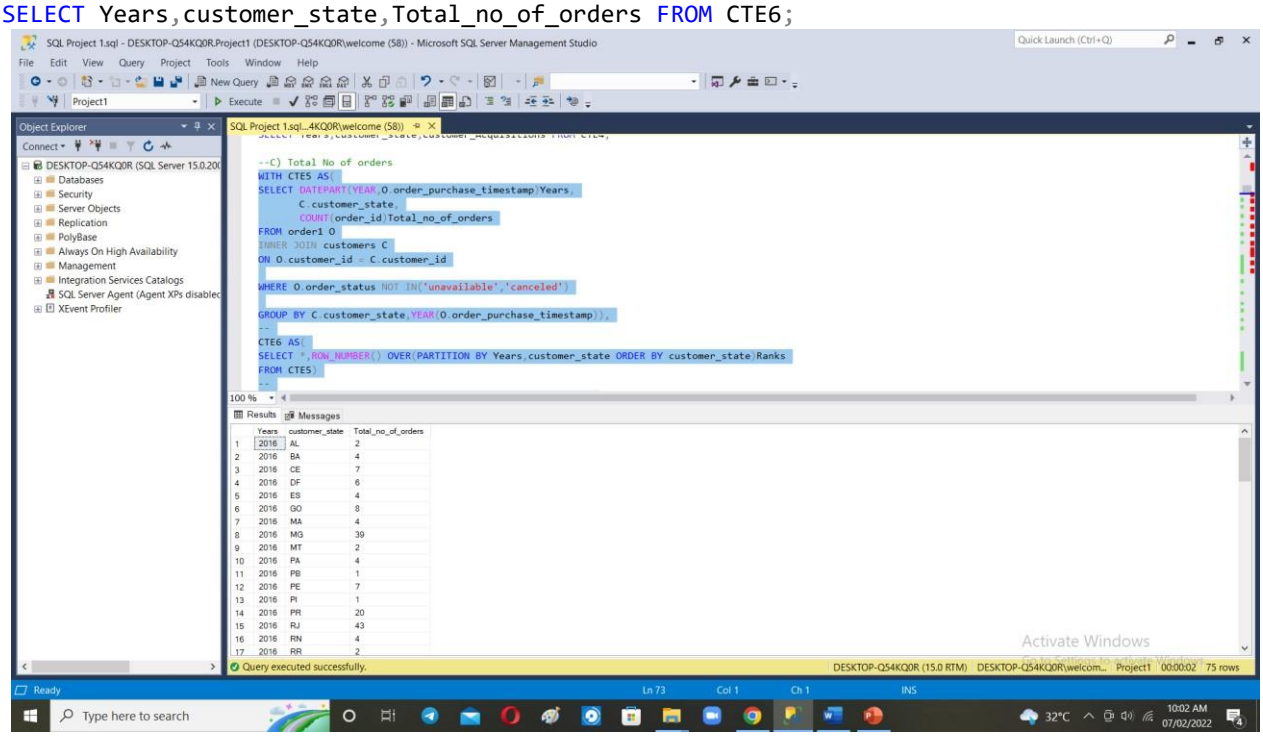

```
--B) Customer Acquisitions
WITH CTE3 AS(
```

```
SELECT
YEAR(O.order_purchase_timestamp)Years,C.customer_state,COUNT(C.customer_id)Customer_Acqui
sitions FROM customers C
INNER JOIN order1 O
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;
```




```
--C) Total No of orders
WITH CTE5 AS(
SELECT DATEPART(YEAR,O.order_purchase_timestamp)Years,
       C.customer_state,
          COUNT(order_id)Total_no_of_orders
FROM order1 O
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.

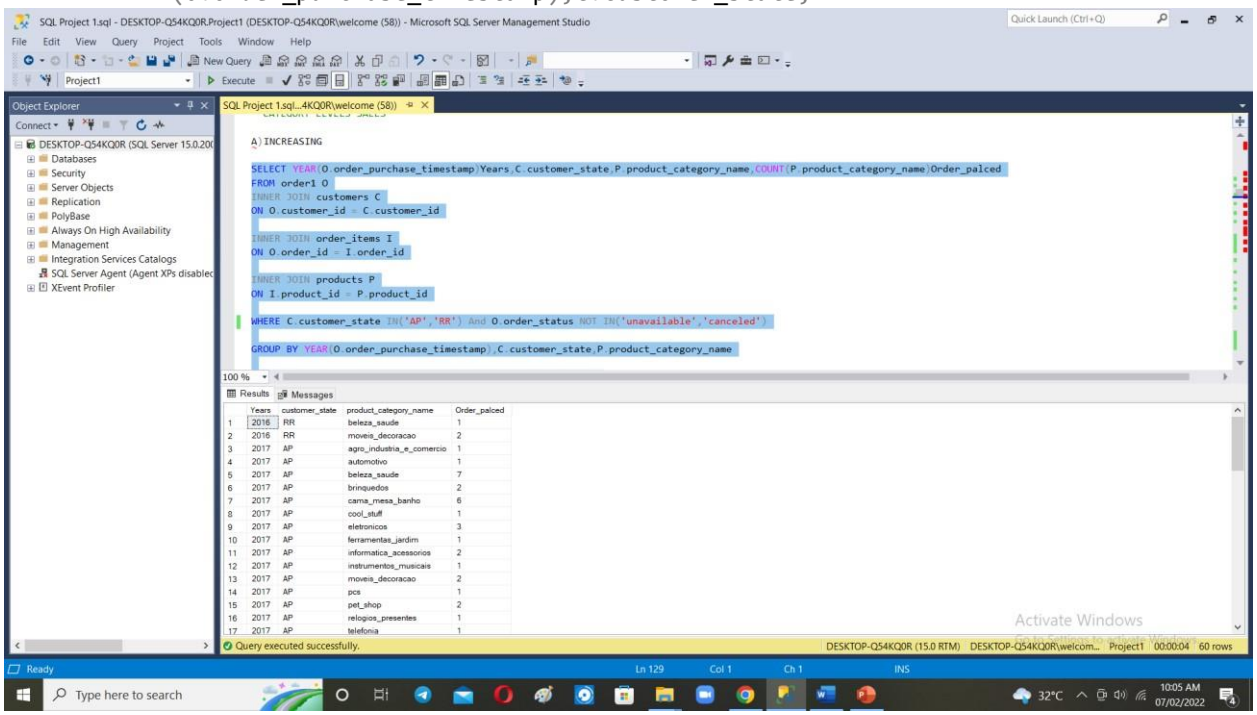
i)Category level Sales and orders placed,

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(O.order_purchase_timestamp),C.customer_state,P.product_category_name
```

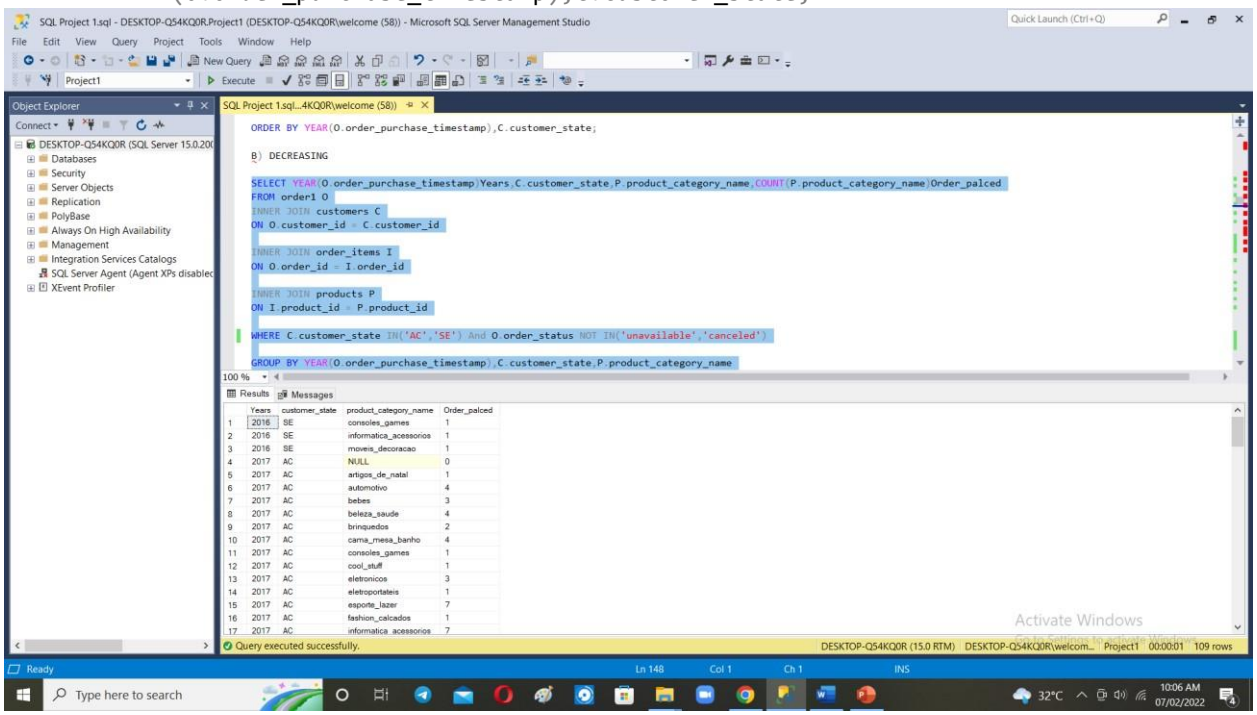
```
ORDER BY YEAR(O.order_purchase_timestamp),C.customer_state;
```



B) DECREASING

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

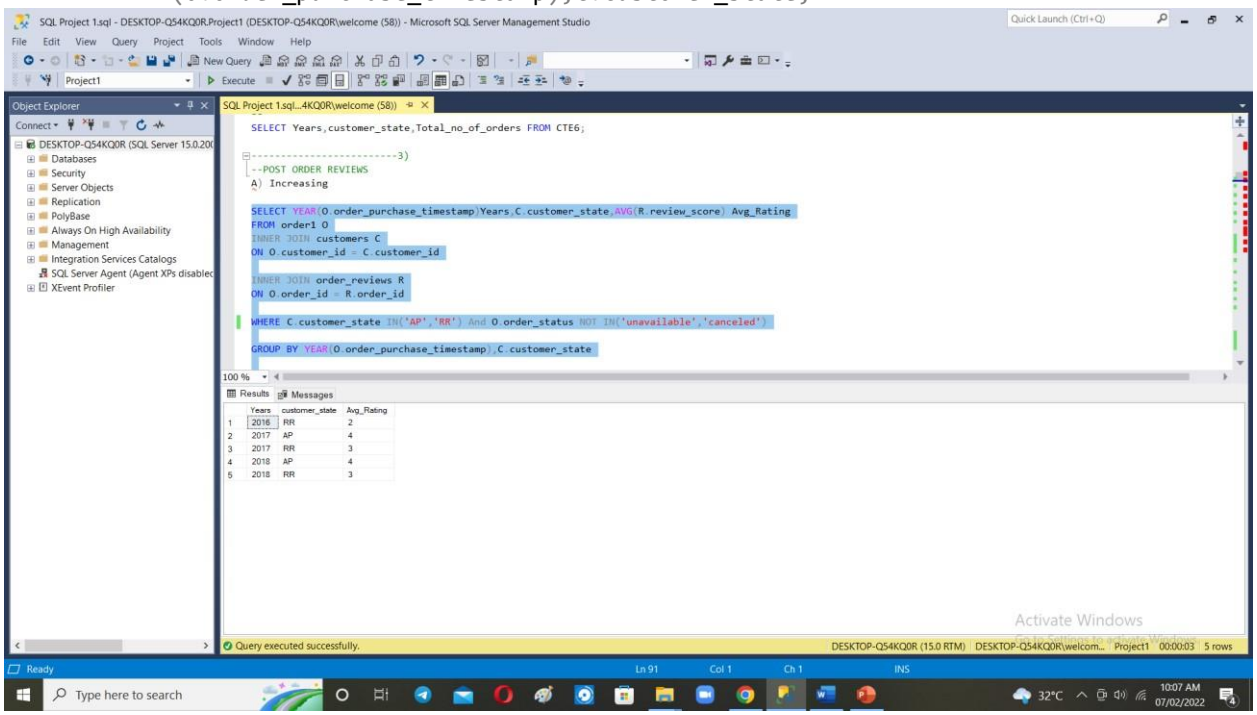
```
ORDER BY YEAR(O.order_purchase_timestamp),C.customer_state;
```



2)post-order reviews

```
A) Increasing

SELECT YEAR(O.order_purchase_timestamp)Years,C.customer_state,AVG(R.review_score)
Avg_Rating
FROM order1 O
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('AP','RR') And O.order_status NOT IN('unavailable','canceled')
GROUP BY YEAR(O.order_purchase_timestamp),C.customer_state
```

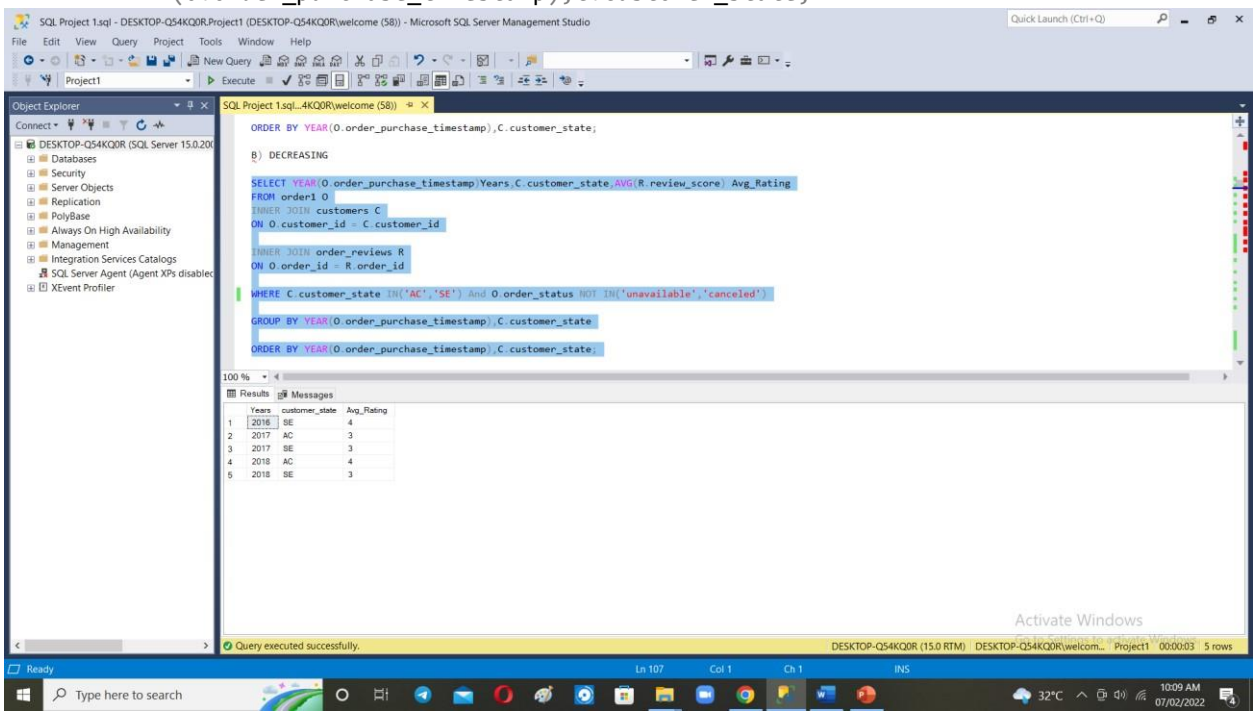
```
ORDER BY YEAR(O.order_purchase_timestamp),C.customer_state;
```



B) DECREASING

```
SELECT YEAR(O.order_purchase_timestamp)Years,C.customer_state,AVG(R.review_score)
Avg_Rating
FROM order1 O
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
```

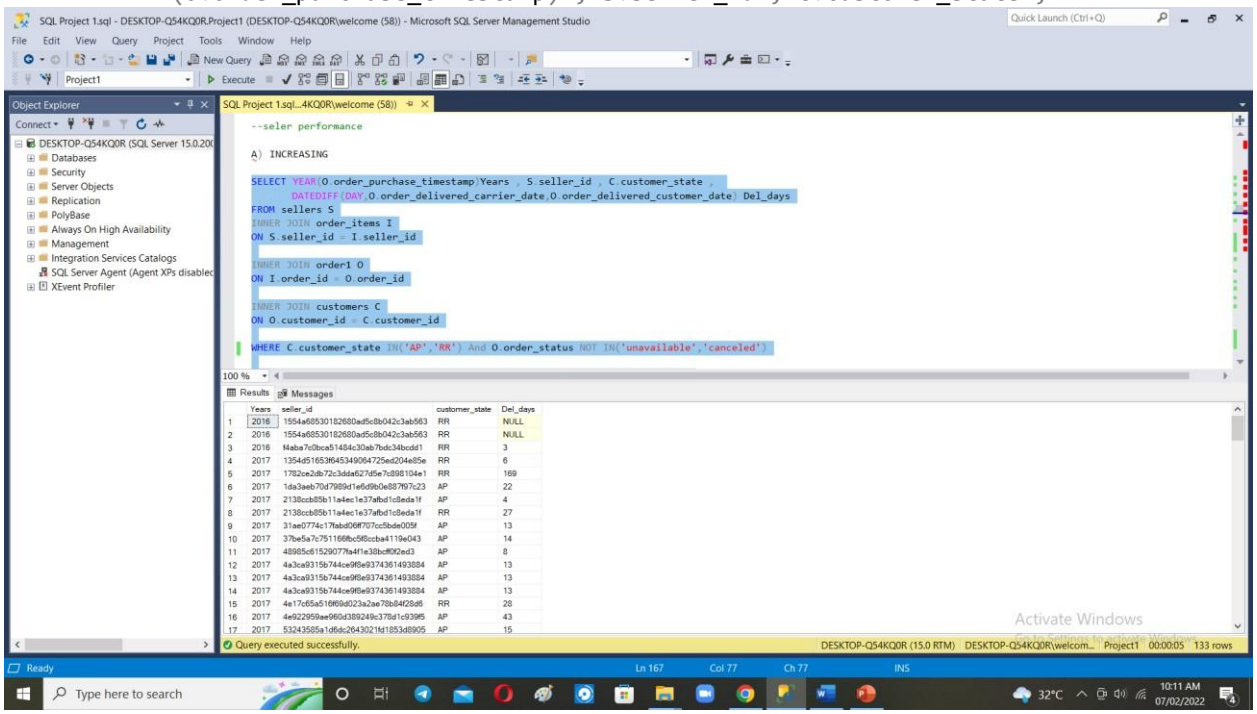
```
ORDER BY YEAR(O.order_purchase_timestamp),C.customer_state;
```



3)Seller performance in terms of deliveries,

```
A) INCREASING

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

INNER JOIN order1 O
ON I.order_id = O.order_id

INNER JOIN customers C
ON O.customer_id = C.customer_id

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

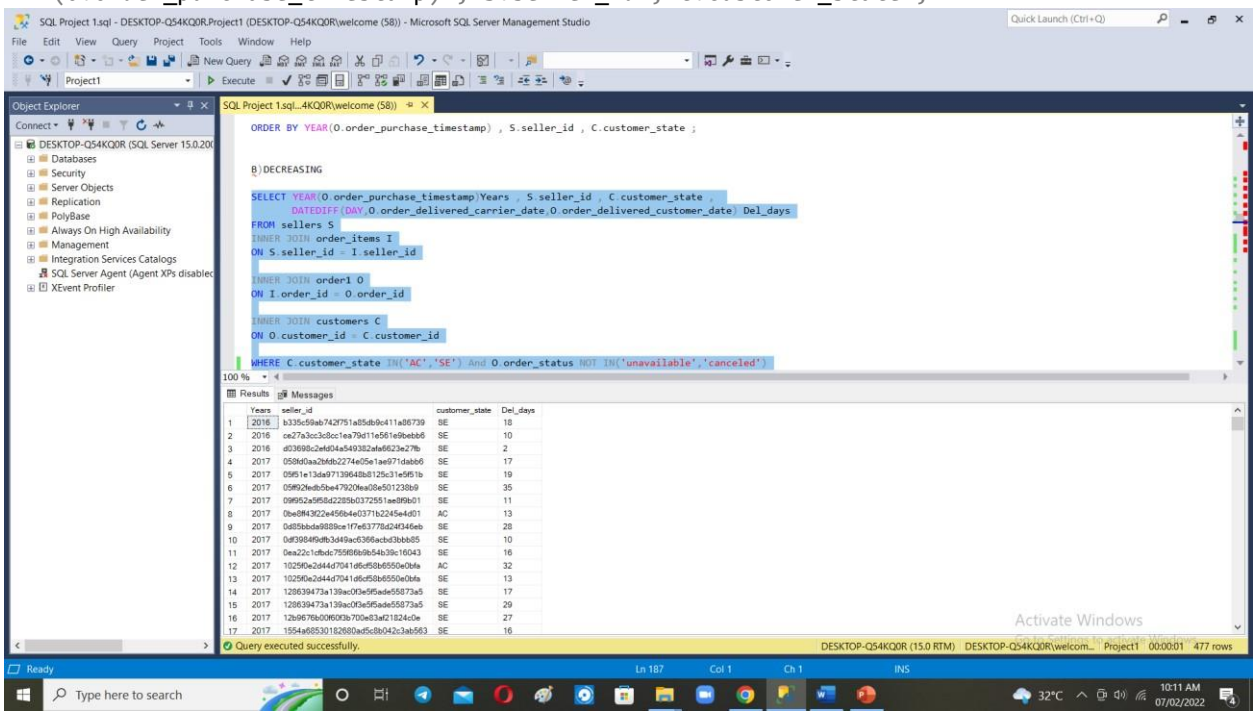
```
ORDER BY
        YEAR(O.order_purchase_timestamp) , S.seller_id , C.customer_state ;
```



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

INNER JOIN order1 O
ON I.order_id = O.order_id

INNER JOIN customers C
ON O.customer_id = C.customer_id
```

```
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 ;
```

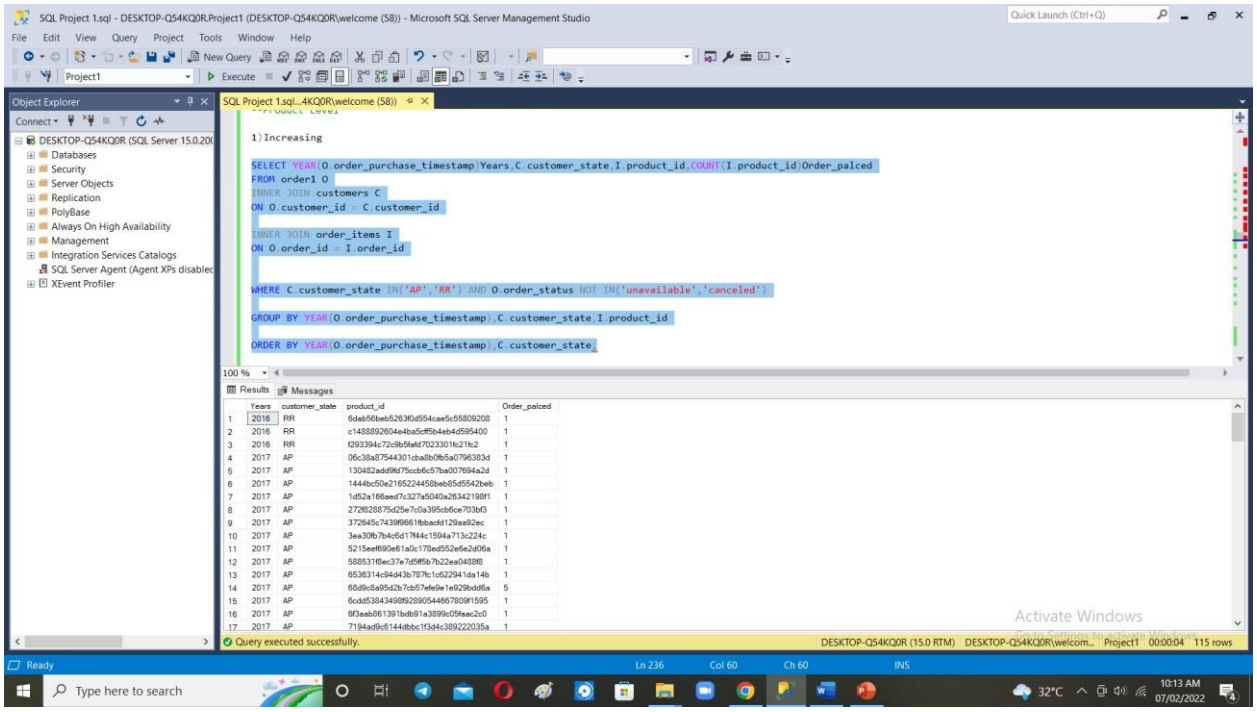


4)product-level sales & orders placed,

1)Increasing

```
SELECT
YEAR(O.order_purchase_timestamp)Years,C.customer_state,I.product_id,COUNT(I.product_id)Or
der_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


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,I.product_id
```

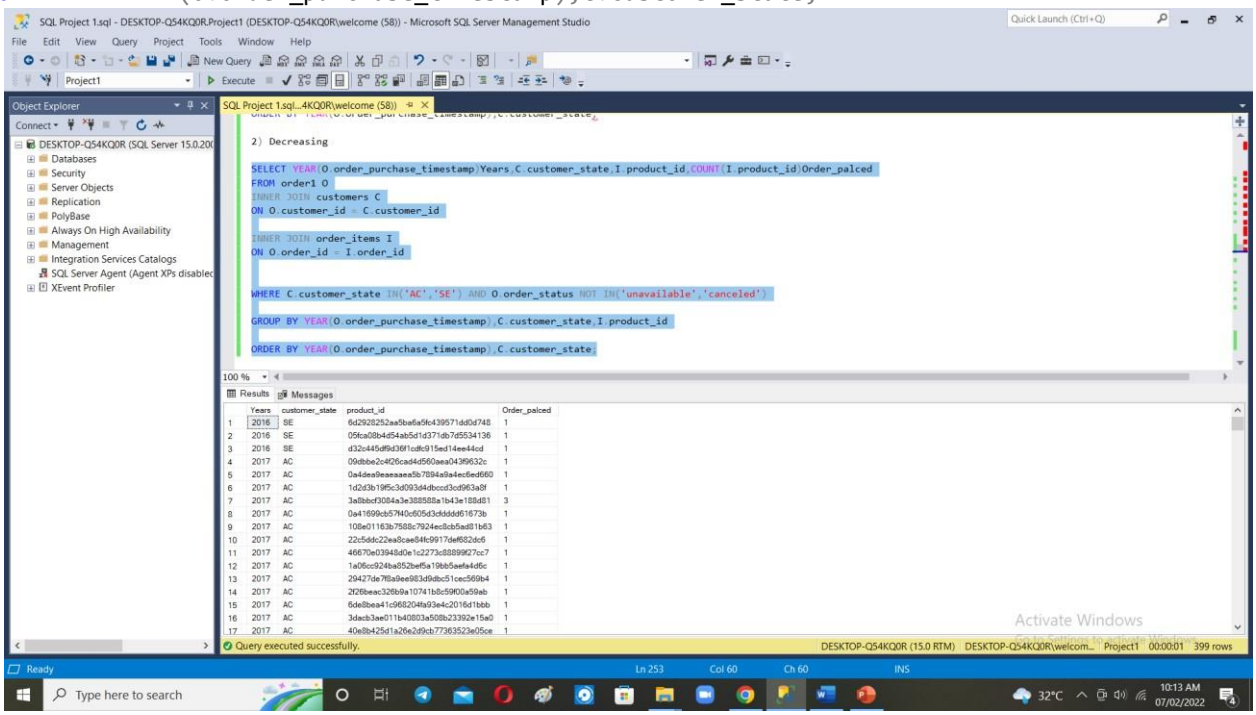
ORDER BY



2) Decreasing

```
SELECT
YEAR(O.order_purchase_timestamp)Years,C.customer_state,I.product_id,COUNT(I.product_id)Or
der_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


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,I.product_id
```

```
ORDER BY YEAR(O.order_purchase_timestamp),C.customer_state;
```

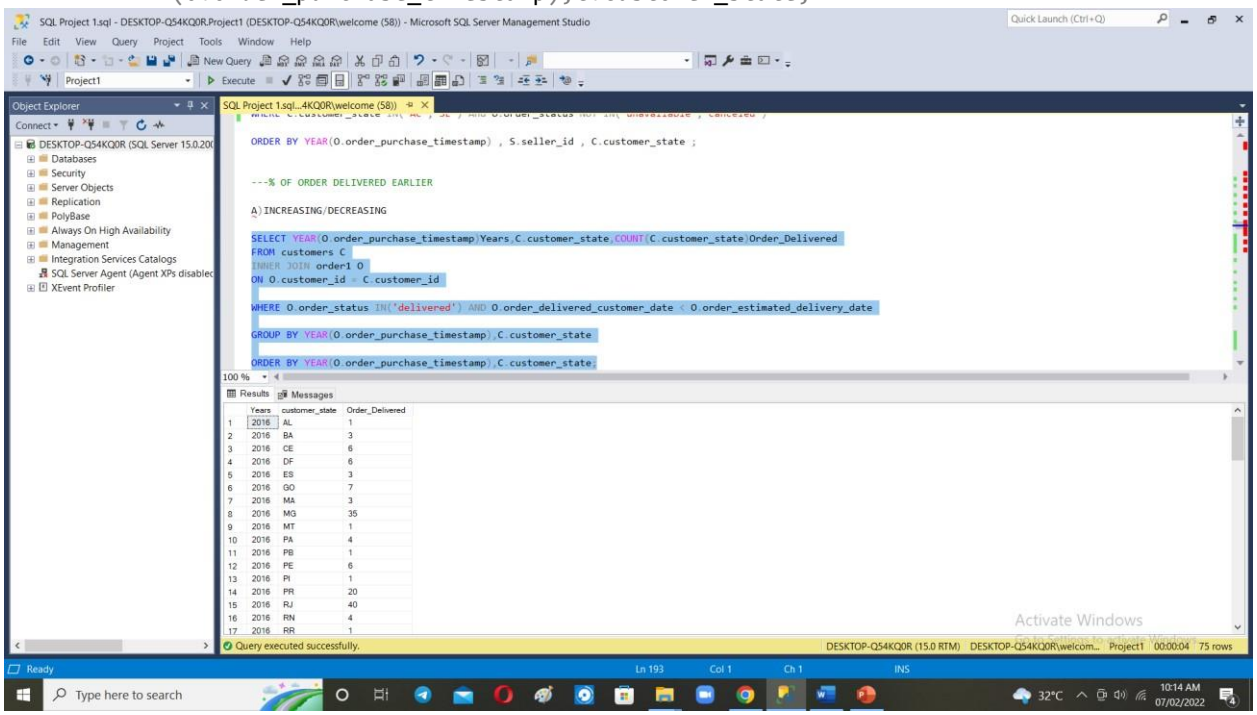


5)% of orders delivered earlier than the expected date,

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 O
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
```

```
ORDER BY YEAR(O.order_purchase_timestamp),C.customer_state;
```

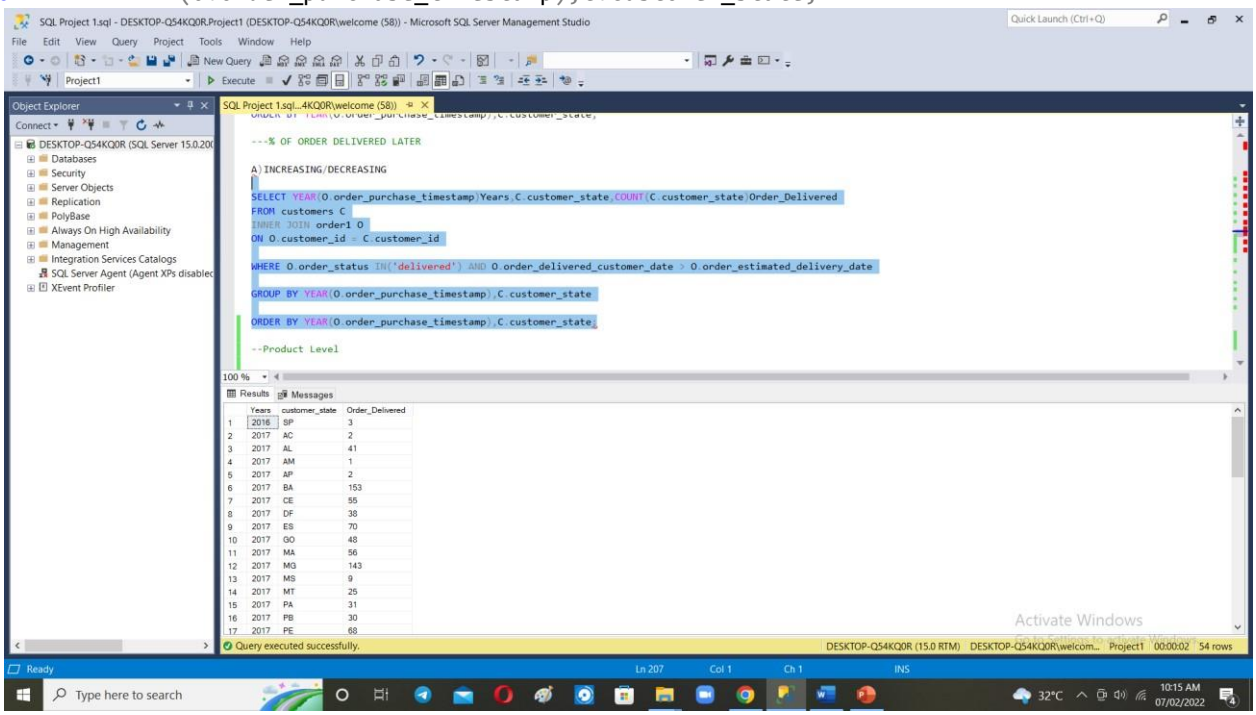


6)% of orders delivered later than the expected date, etc.e

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 O
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
```

```
ORDER BY YEAR(O.order_purchase_timestamp),C.customer_state;
```



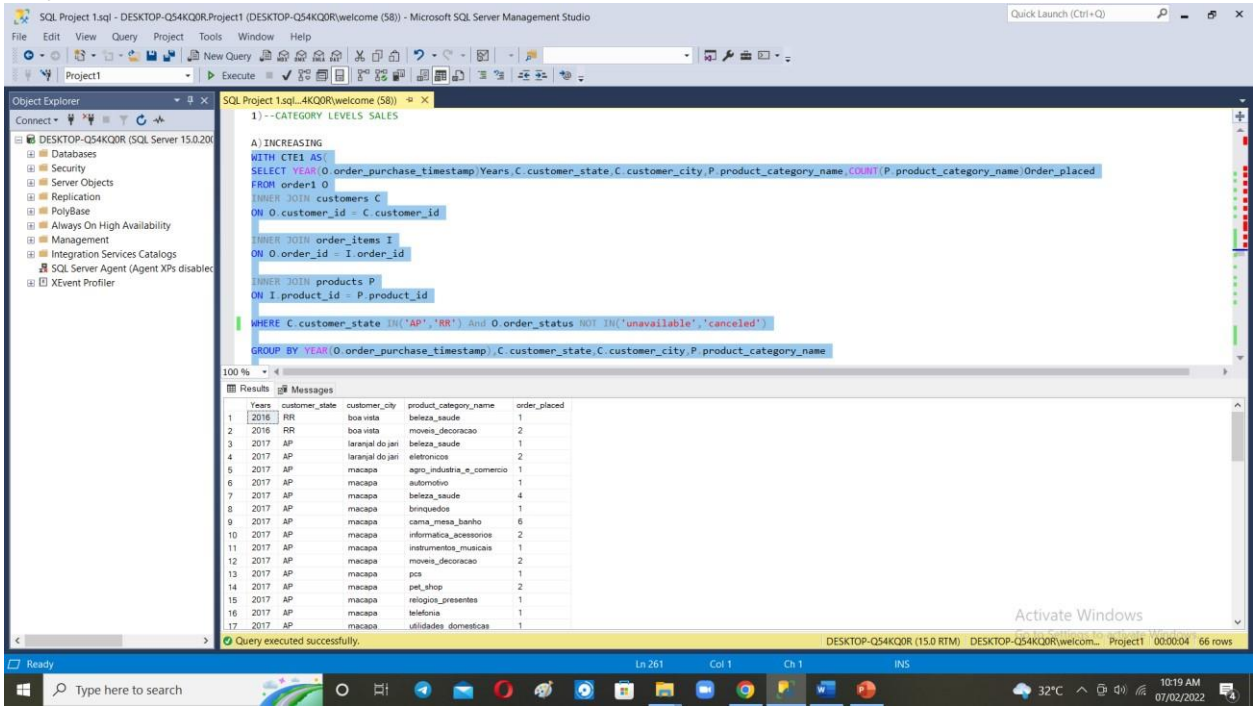
D. Do the above analysis for the top 2 cities which are causing the trend for each of the states identified in point (b)

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 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(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
```

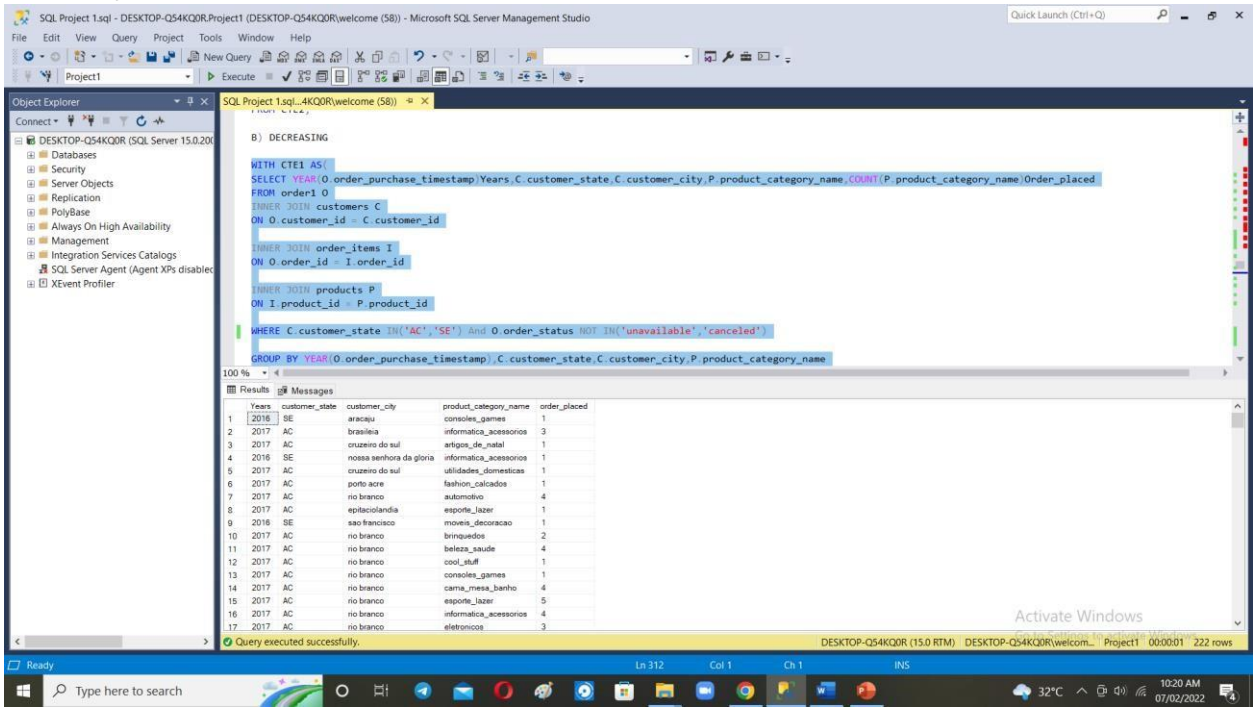
```
ORDER BY YEAR(O.order_purchase_timestamp),C.customer_state;
FROM CTE1)
```

```
--
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 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('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;
```

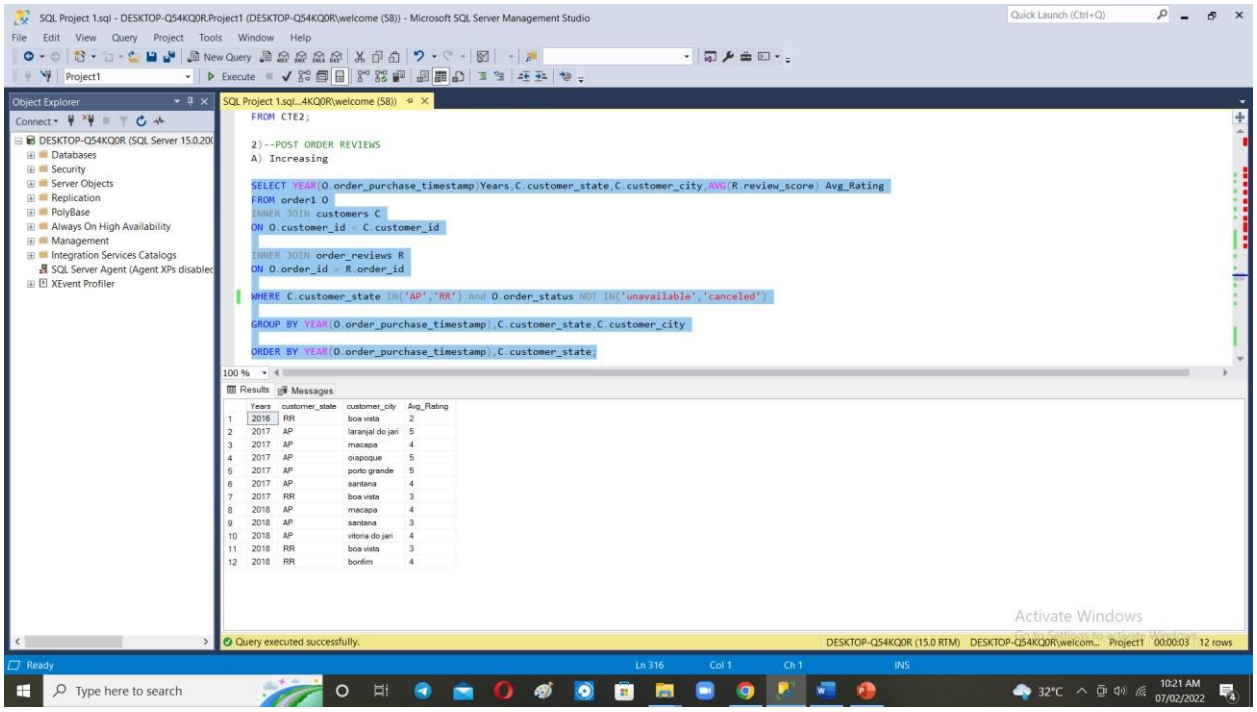


2) post-order reviews,

A) Increasing

```
SELECT
YEAR(O.order_purchase_timestamp)Years,C.customer_state,C.customer_city,AVG(R.review_score
) Avg_Rating
FROM order1 O
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('AP','RR') And O.order_status NOT IN('unavailable','canceled')
GROUP BY YEAR(O.order_purchase_timestamp),C.customer_state,C.customer_city
YEAR(O.order_purchase_timestamp),C.customer_state;
```

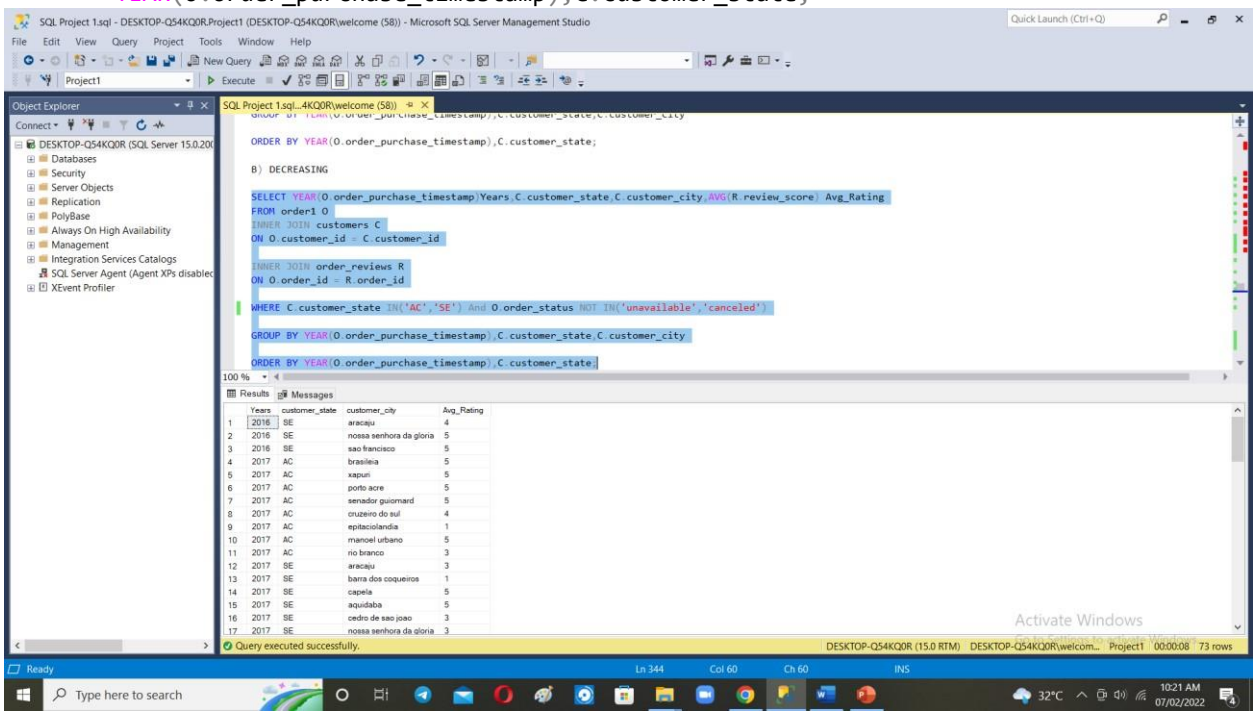
```
ORDER BY
```



B) DECREASING

```
SELECT
YEAR(O.order_purchase_timestamp)Years,C.customer_state,C.customer_city,AVG(R.review_score
) Avg_Rating
FROM order1 O
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
```

```
ORDER BY
        YEAR(O.order_purchase_timestamp),C.customer_state;
```

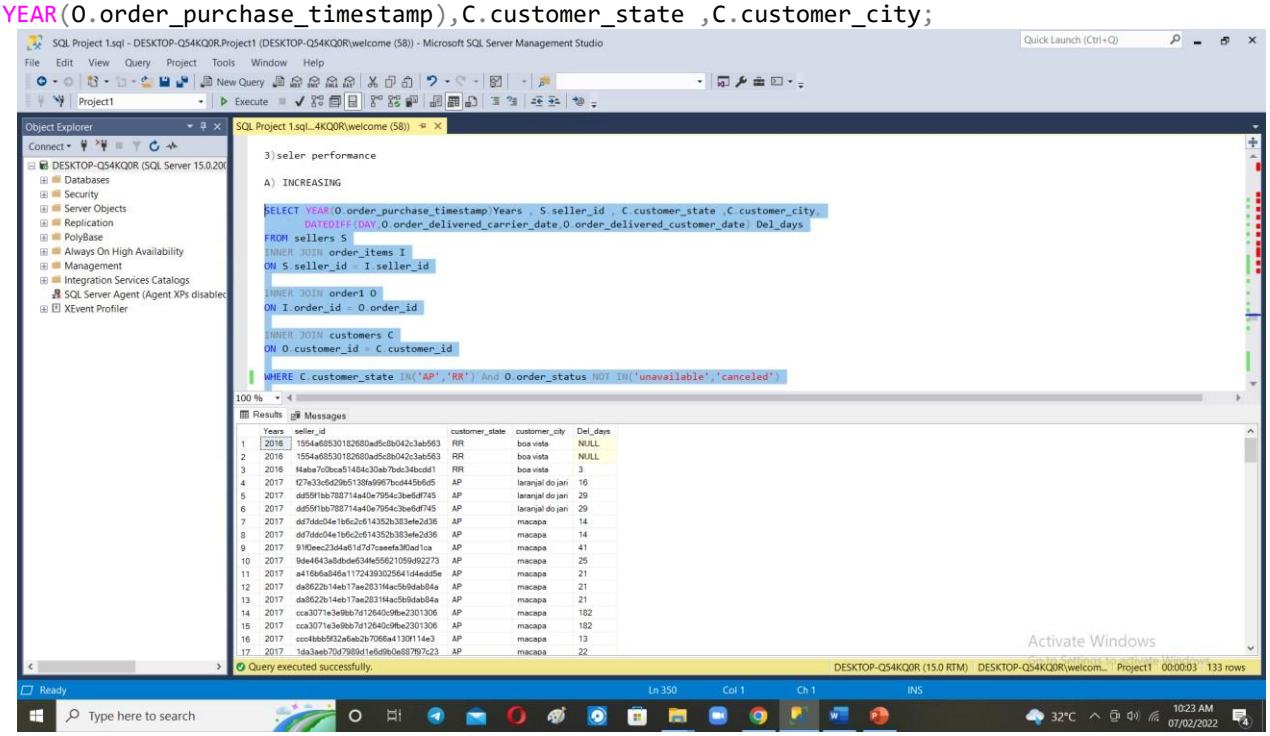


3)Seller performance in terms of deliveries,

A) INCREASING

```
SELECT YEAR(O.order_purchase_timestamp)Years , S.seller_id , C.customer_state
,C.customer_city,
       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

INNER JOIN order1 O
ON I.order_id = O.order_id

INNER JOIN customers C
```

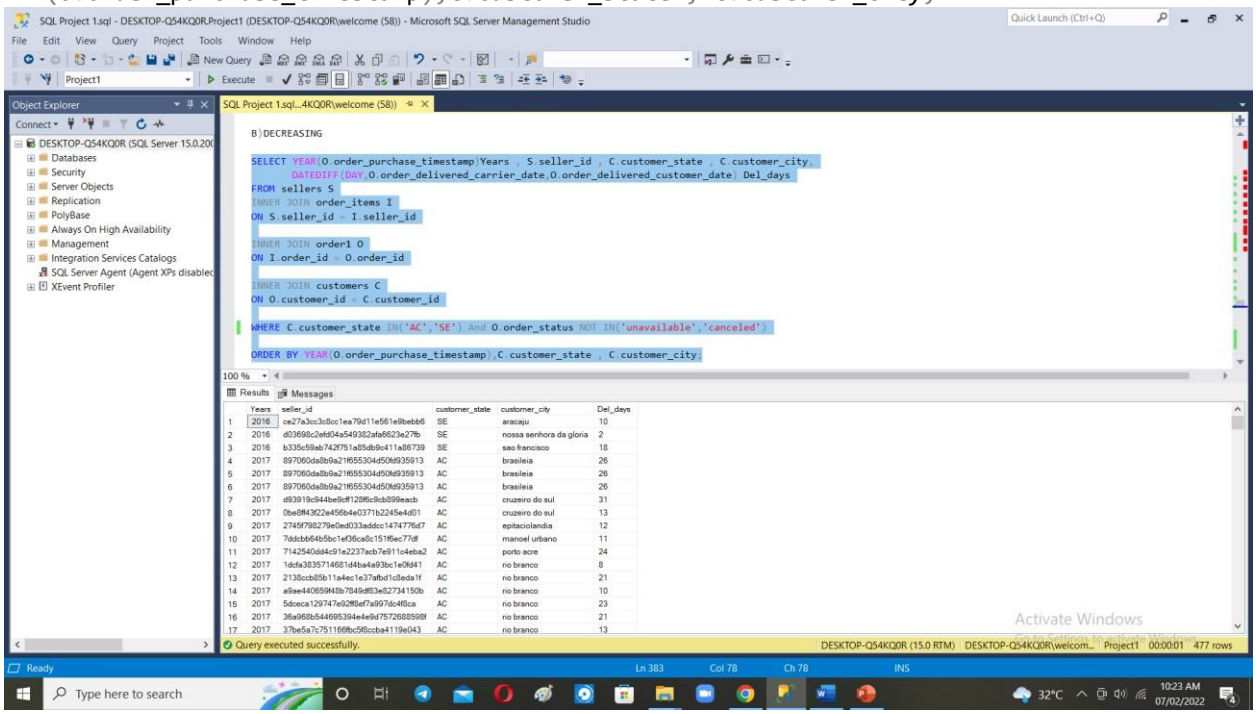
```
ORDER BY
ON O.customer_id = C.customer_id

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

```
SELECT YEAR(O.order_purchase_timestamp)Years , S.seller_id , C.customer_state ,
C.customer_city,
      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

INNER JOIN order1 O
```

```
ORDER BY
ON I.order_id = O.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;
```



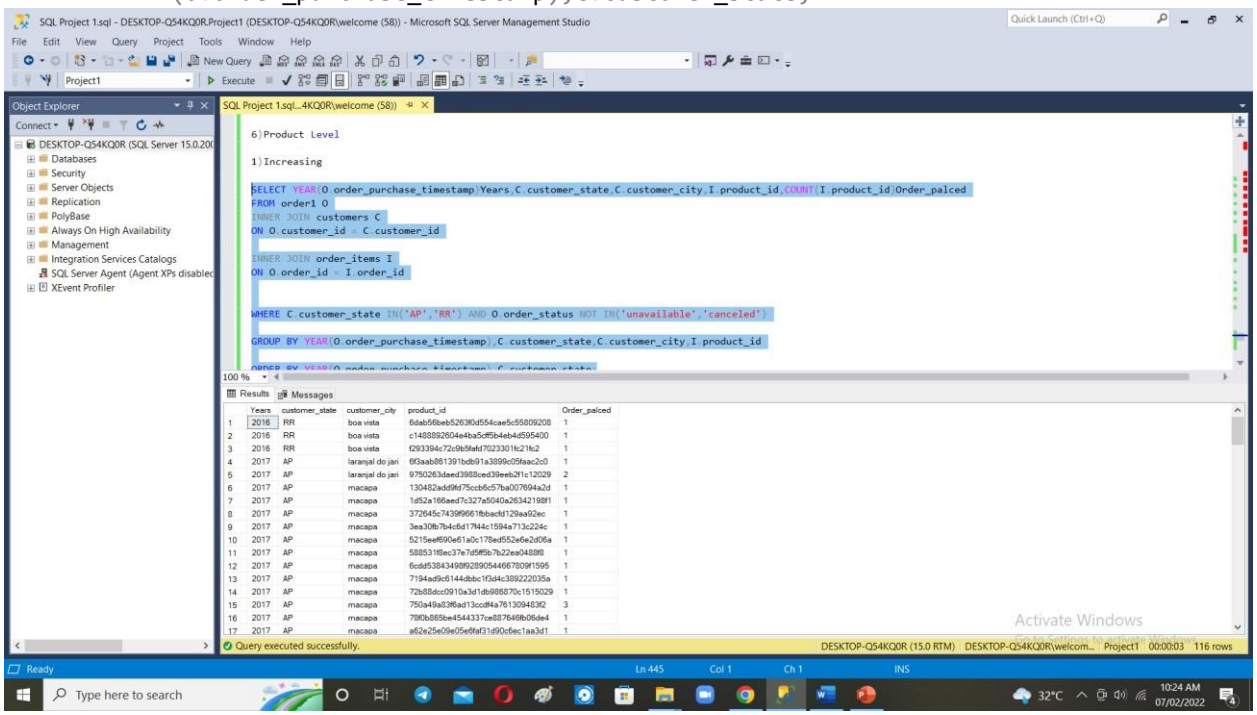
4)product-level sales & orders placed,

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 O
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
```

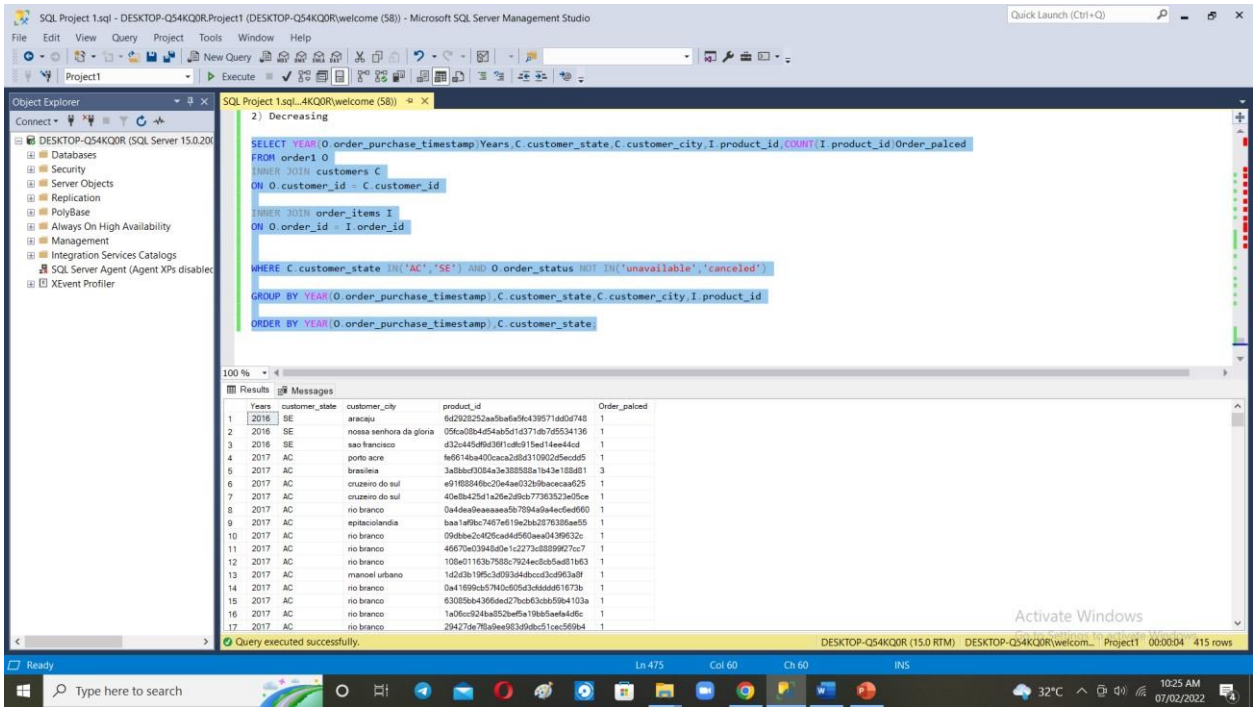
```
ORDER BY YEAR(O.order_purchase_timestamp),C.customer_state;
```



2) Decreasing

```
SELECT
YEAR(O.order_purchase_timestamp)Years,C.customer_state,C.customer_city,I.product_id,COUNT
(I.product_id)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


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,I.product_id
```

```
ORDER BY YEAR(O.order_purchase_timestamp),C.customer_state;
```

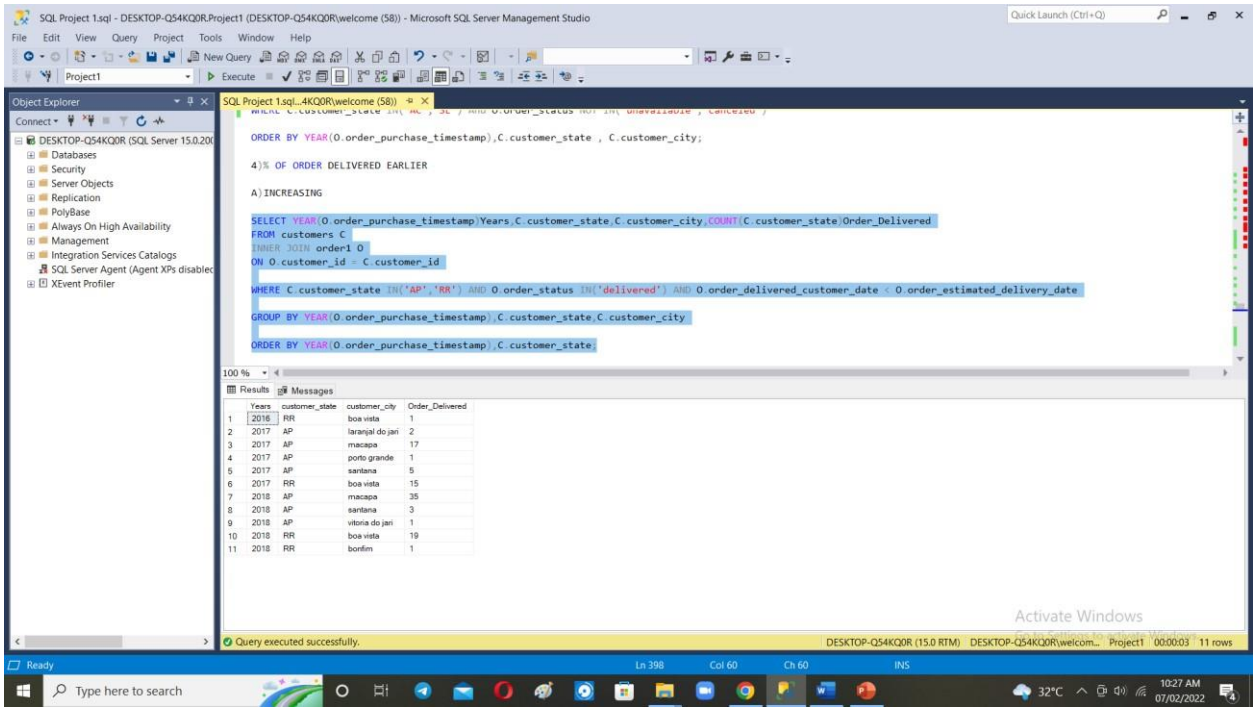


5)% of orders delivered earlier than the expected date,

A)INCREASING

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

WHERE C.customer_state IN('AP','RR') 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
```

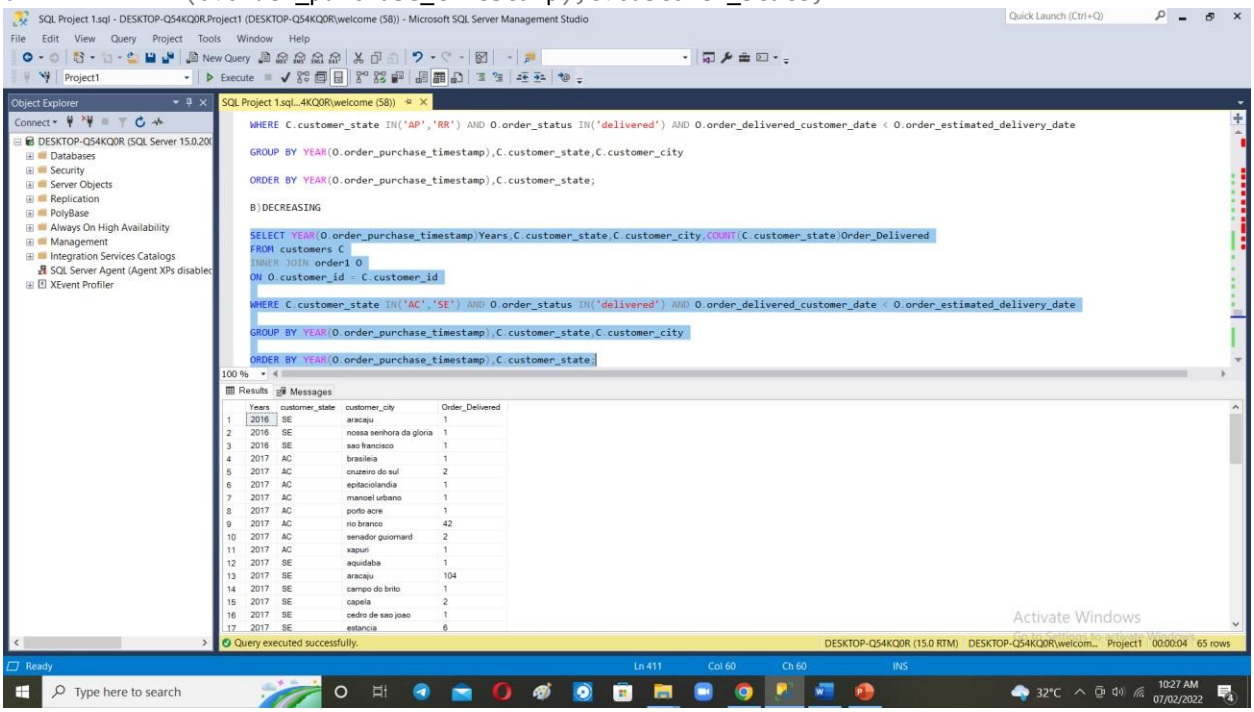
```
ORDER BY YEAR(O.order_purchase_timestamp),C.customer_state;
```



B)DECREASING

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

```
ORDER BY YEAR(O.order_purchase_timestamp),C.customer_state;
```

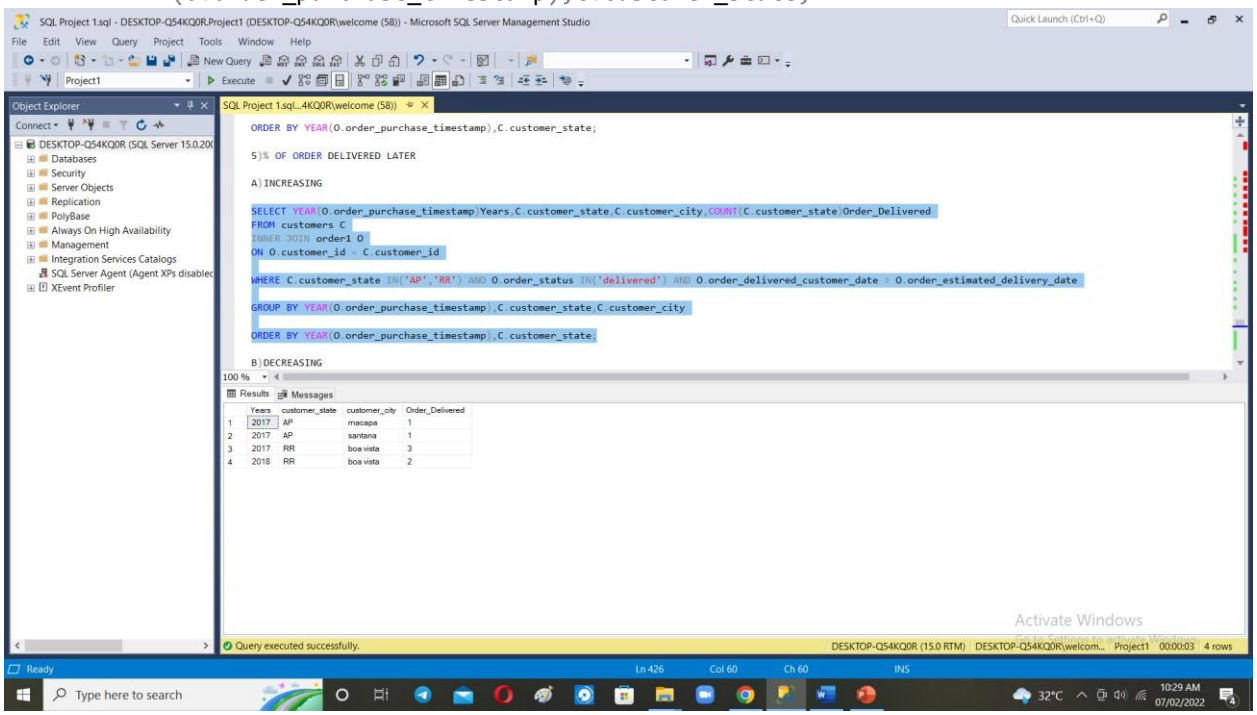


6)% of orders delivered later than the expected date, etc.e

A)INCREASING

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

WHERE C.customer_state IN('AP','RR') 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
```

```
ORDER BY YEAR(O.order_purchase_timestamp),C.customer_state;
```



B)DECREASING

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

```
ORDER BY YEAR(O.order_purchase_timestamp),C.customer_state;
```