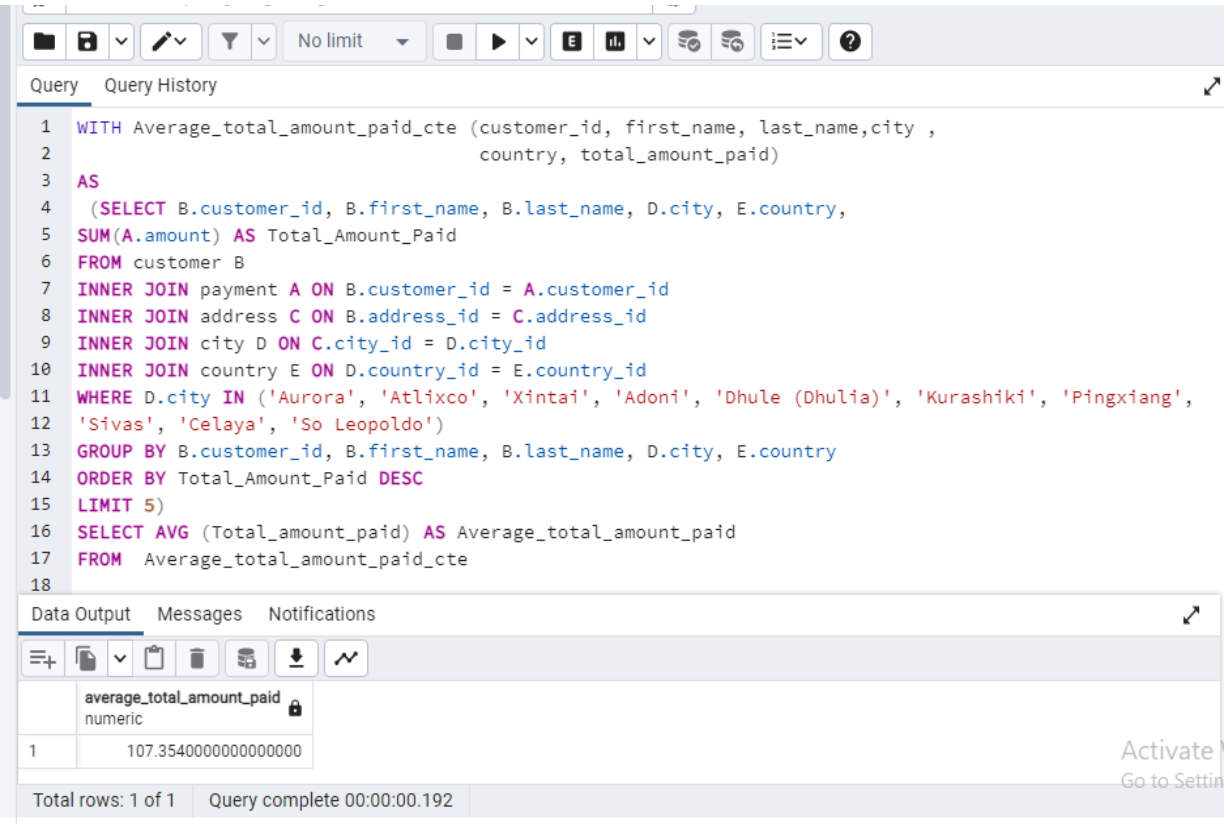
3.9 Common Table Expression(CTE)

Step 1: Answer the business questions from step 1 and 2 of task 3.8 using CTEs

1. Rewrite your queries from steps 1 and 2 of task 3.8 as CTEs.

- Find the average amount paid by the top 5 customers

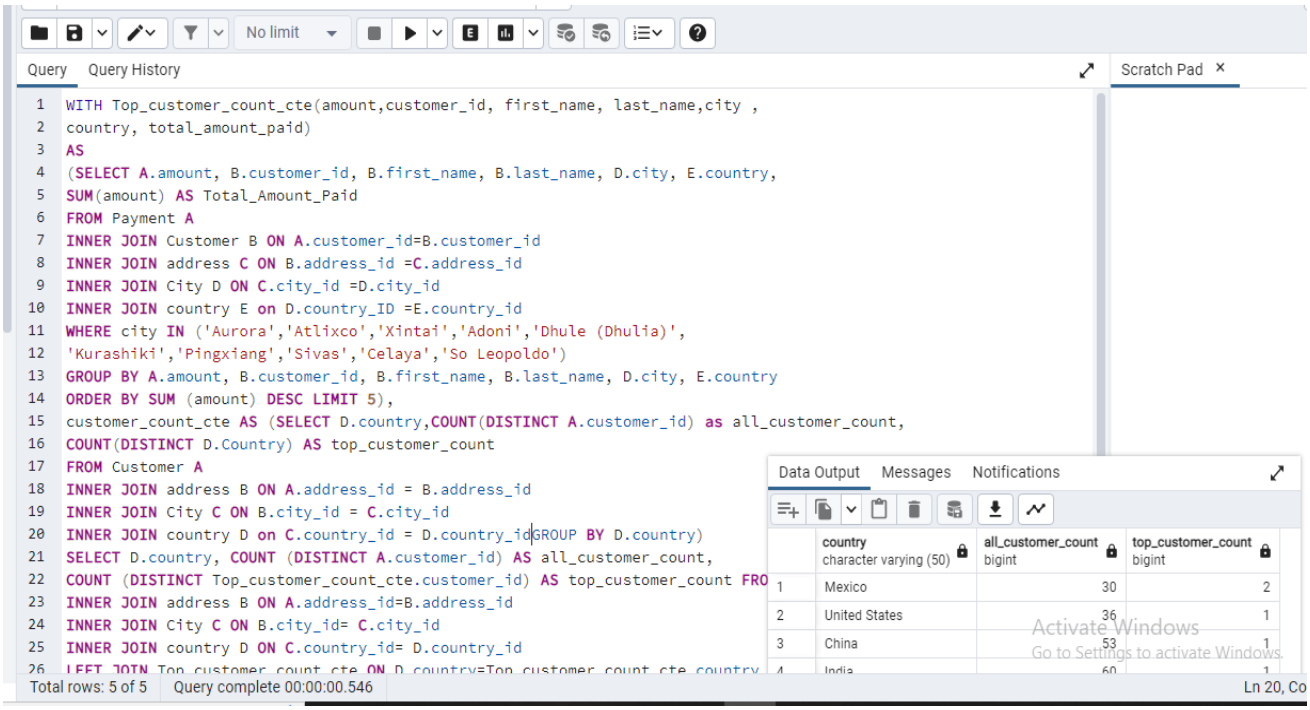
```
WITH Average_total_amount_paid_cte (customer_id, first_name, last_name,city ,
country, total_amount_paid)
AS
 (SELECT B.customer_id, B.first_name, B.last_name, D.city, E.country,
SUM(A.amount) AS Total_Amount_Paid
FROM customer B
INNER JOIN payment A ON B.customer_id = A.customer_id
INNER JOIN address C ON B.address_id = C.address_id
INNER JOIN city D ON C.city_id = D.city_id
INNER JOIN country E ON D.country_id = E.country_id
WHERE D.city IN ('Aurora', 'Atlixco', 'Xintai', 'Adoni', 'Dhule (Dhulia)', 'Kurashiki', 'Pingxiang',
'Sivas', 'Celaya', 'So Leopoldo')
GROUP BY B.customer_id, B.first_name, B.last_name, D.city, E.country
ORDER BY Total_Amount_Paid DESC
LIMIT 5)
SELECT AVG (Total_amount_paid) AS Average_total_amount_paid
FROM  Average_total_amount_paid_cte
```



STEP 2:

• Find the out how many of the top 5 customers are based within each country.

```
WITH Top_customer_count_cte(amount,customer_id, first_name, last_name,city ,
country, total_amount_paid)
AS
(SELECT A.amount, B.customer_id, B.first_name, B.last_name, D.city, E.country,
SUM(amount) AS Total_Amount_Paid
FROM Payment A
INNER JOIN Customer B ON A.customer_id=B.customer_id
INNER JOIN address C ON B.address_id =C.address_id
INNER JOIN City D ON C.city_id =D.city_id
INNER JOIN country E on D.country_ID =E.country_id
WHERE city IN ('Aurora','Atlixco','Xintai','Adoni','Dhule (Dhulia)',
'Kurashiki','Pingxiang','Sivas','Celaya','So Leopoldo')
GROUP BY A.amount, B.customer_id, B.first_name, B.last_name, D.city, E.country
ORDER BY SUM (amount) DESC LIMIT 5),
customer_count_cte AS (SELECT D.country,COUNT(DISTINCT A.customer_id) as
all_customer_count,
COUNT(DISTINCT D.Country) AS top_customer_count
FROM Customer A
INNER JOIN address B ON A.address_id = B.address_id
INNER JOIN City C ON B.city_id = C.city_id
INNER JOIN country D on C.country_id = D.country_id
GROUP BY D.country)
SELECT D.country, COUNT (DISTINCT A.customer_id) AS all_customer_count,
COUNT (DISTINCT Top_customer_count_cte.customer_id) AS top_customer_count FROM
Customer A
INNER JOIN address B ON A.address_id=B.address_id
INNER JOIN City C ON B.city_id= C.city_id
INNER JOIN country D ON C.country_id= D.country_id
LEFT JOIN Top_customer_count_cte ON D.country=Top_customer_count_cte.country
GROUP BY D.Country
ORDER BY top_customer_count DESC
LIMIT 5
```



2. Write 2 to 3 sentences explaining how you approached this step, for example, what you did first,second, and so on.

For step1, I copied the syntax from 3.8 assignment , then I removed outer query from the subquery statement and replaced it with CTE. Using the WITH Clause and followed by AS keyword and at the end write, the Statement to find the average of the Top 5 Paying customer in the top 10 city.

For step 2, I used two times CTE, one for the count of customer and other for count of the top customer in countries.

Q2: Compare the performance of your CTEs and subqueries.

- Which approach do you think will perform better and why?
  I found diificult in writing CTE queries and Sub query found easy to write.
  But ,after cost estimation and speed, I found CTE is best approach .

- Compare the costs of all the queries by creating query plans for each one.

Step 1	Subquery	CTE
Cost	cost=65.06..65.07 rows=1 width=32	cost=65.06. 65.07 rows=1 width=32
Time	Total query runtime: 298 msec. 22 rows affected.	Total query runtime: 267 msec. 22 rows affected.
Step 2		
Cost	cost=180.55..180.56 rows=5 width=90	cost=166.83. 166.85 rows=5 width=25
Time	Total query runtime: 262 msec. 45 rows affected	Total query runtime: 157 msec. 46 rows affected.

- The EXPLAIN command gives you an estimated cost. To find out the actual speed of your queries, run them in pgAdmin 4. After each query has been run, a pop-up window will display its speed in milliseconds.
- Did the results surprise you? Write a few sentences to explain your answer
  For Step 1, the cost is the same ,but CTE is faster than subquery. In step 2,the cost and time is less in both CTE and Subquery.
  I think opting for theoption CTE is best option than Subquery.

Q3: Step 3: Write 1 to 2 paragraphs on the challenges you faced when replacing your subqueries with CTEs

In performing the task ,I faced a lot of problems. Task 1 was quites easy as compare to task2. The result was not coming accurately and when I look at example of other,I got to realize I forget to put Cte in second Cte statement and it was always showing error and than I correct it. Got result. But I personally found it more diificult and time consuming as compare to subqueries.