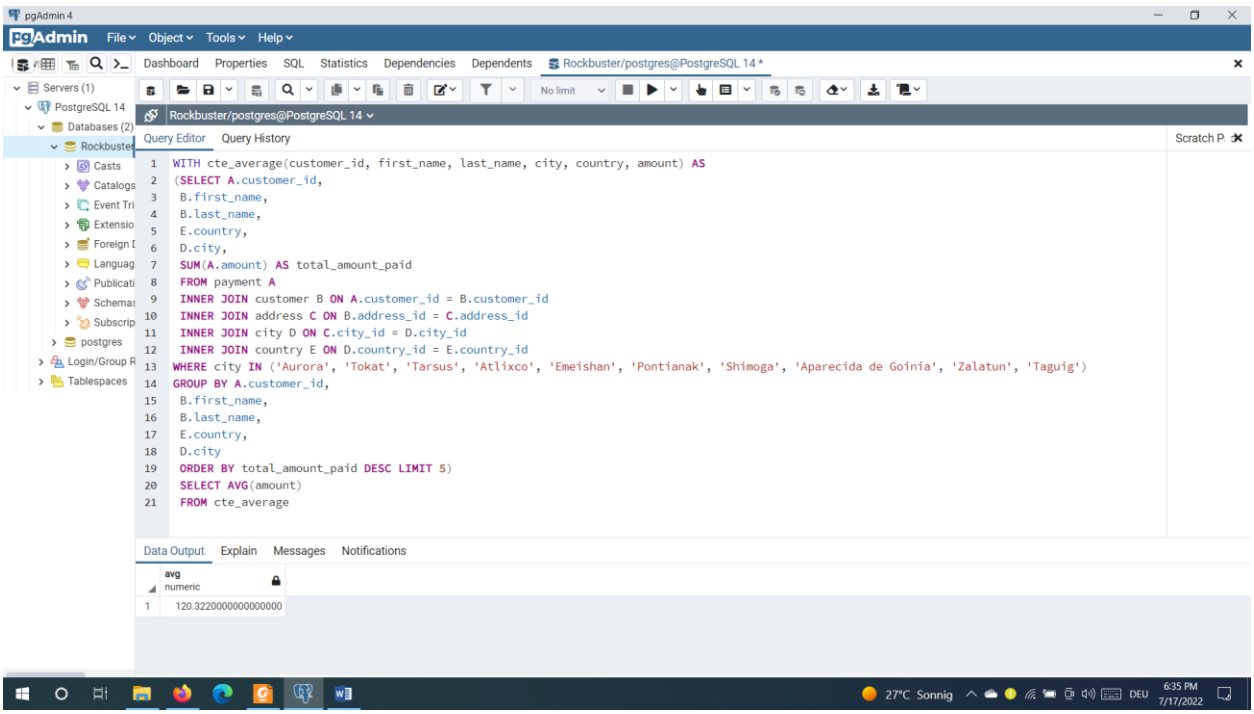
3.9 Common Table Expressions

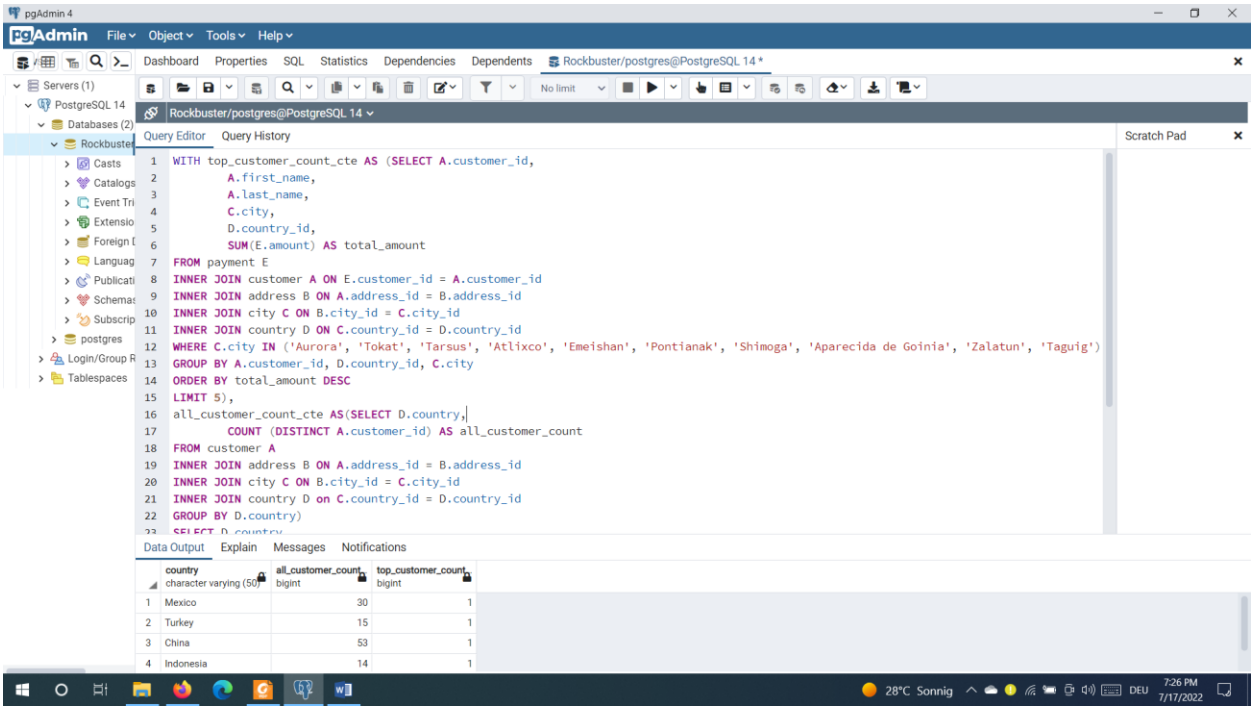
Step 1: Answer the business questions from step 1 and 2 of task 3.8 using CTEs

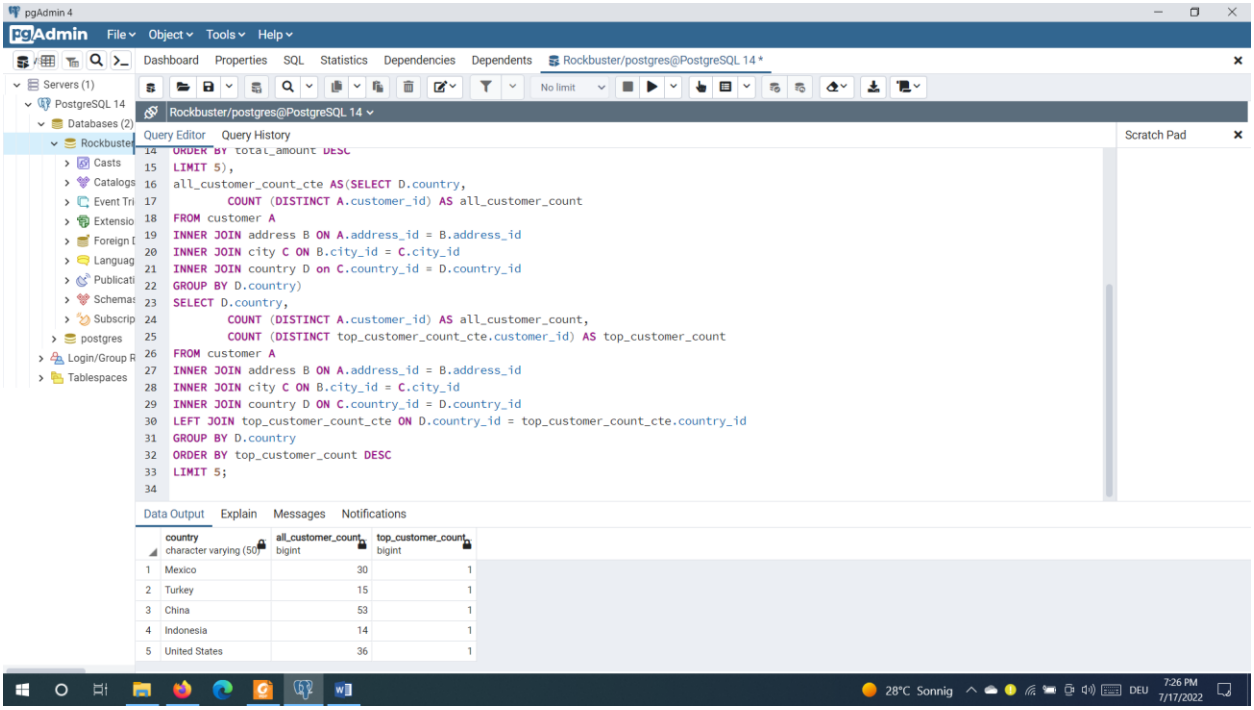
1. Rewrite your queries from steps 1 and 2 of task 3.8 as CTEs.
2. Copy-paste your CTEs and their outputs into your answers document.
3. Write 2 to 3 sentences explaining how you approached this step, for example, what you did first, second, and so on.

1.a)



1.b)


pgAdmin 4
File Object Tools Help
Dashboard Properties SQL Statistics Dependencies Dependents Rockbuster/postgres@PostgreSQL 14 *
Rockbuster/postgres@PostgreSQL 14
Query Editor Query History
Scratch Pad
WITH top_customer_count_cte AS (SELECT A.customer_id,
A.first_name,
A.last_name,
C.city,
D.country_id,
SUM(E.amount) AS total_amount
FROM payment E
INNER JOIN customer A ON E.customer_id = A.customer_id
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
WHERE C.city IN ('Aurora', 'Tokat', 'Tarsus', 'Atlixco', 'Emeishan', 'Pontianak', 'Shimoga', 'Aparecida de Goinia', 'Zalatun', 'Taguig')
GROUP BY A.customer_id, D.country_id, C.city
ORDER BY total_amount DESC
LIMIT 5),
all_customer_count_cte AS(SELECT D.country,
COUNT (DISTINCT A.customer_id) AS all_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)
Data Output Explain Messages Notifications
country character varying (50) all_customer_count bigint top_customer_count bigint
1 Mexico 30 1
2 Turkey 15 1
3 China 53 1
4 Indonesia 14 1
28°C Sonnig DEU 7:26 PM 7/17/2022


pgAdmin 4
File Object Tools Help
Dashboard Properties SQL Statistics Dependencies Dependents Rockbuster/postgres@PostgreSQL 14 *
Rockbuster/postgres@PostgreSQL 14
Query Editor Query History
Scratch Pad
ORDER BY total_amount DESC
LIMIT 5),
all_customer_count_cte AS(SELECT D.country,
COUNT (DISTINCT A.customer_id) AS all_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_id = top_customer_count_cte.country_id
GROUP BY D.country
ORDER BY top_customer_count DESC
LIMIT 5;
Data Output Explain Messages Notifications
country character varying (50) all_customer_count bigint top_customer_count bigint
1 Mexico 30 1
2 Turkey 15 1
3 China 53 1
4 Indonesia 14 1
5 United States 36 1
28°C Sonnig DEU 7:26 PM 7/17/2022

1.c) I have defined the CTE with the 'WITH' command, as well as necessary columns. I have set the name for the CTE to a descriptive name. I have wrote the main statement by inserting already the written queries from task 3.8 in parentheses after the AS command. I have adjusted order and grouping accordingly.

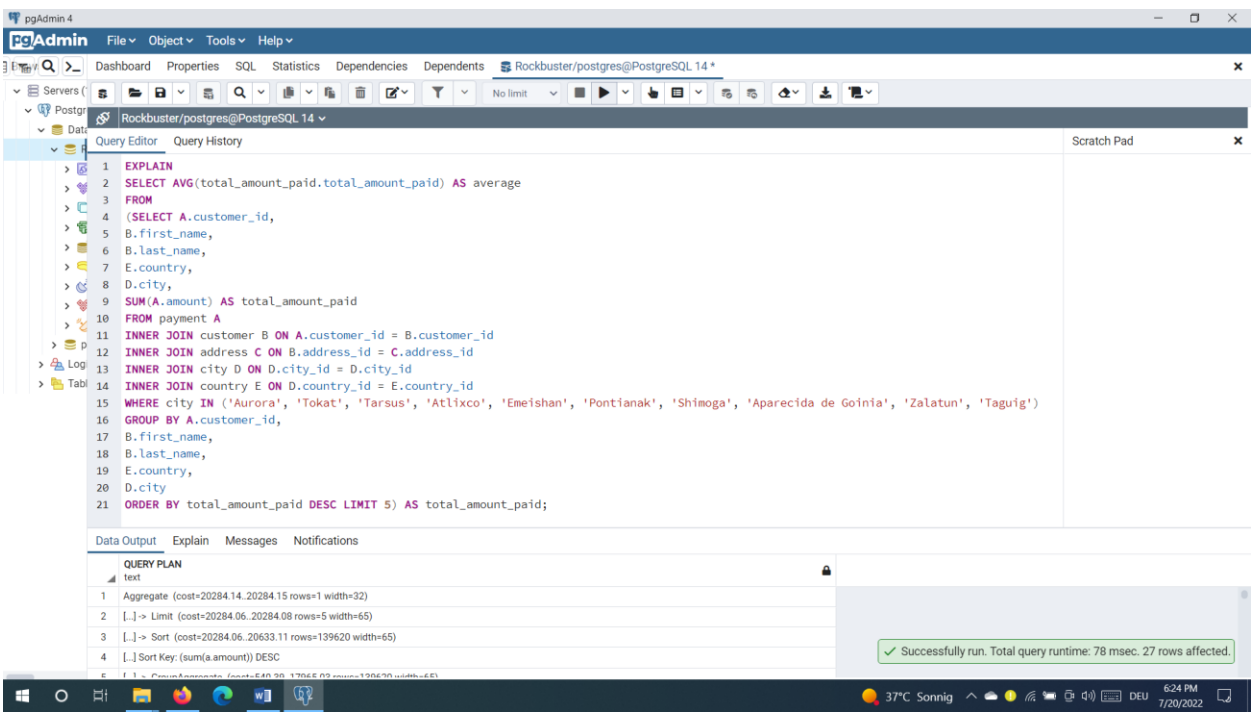
Step 2: Compare the performance of your CTEs and subqueries.

1. Which approach do you think will perform better and why?

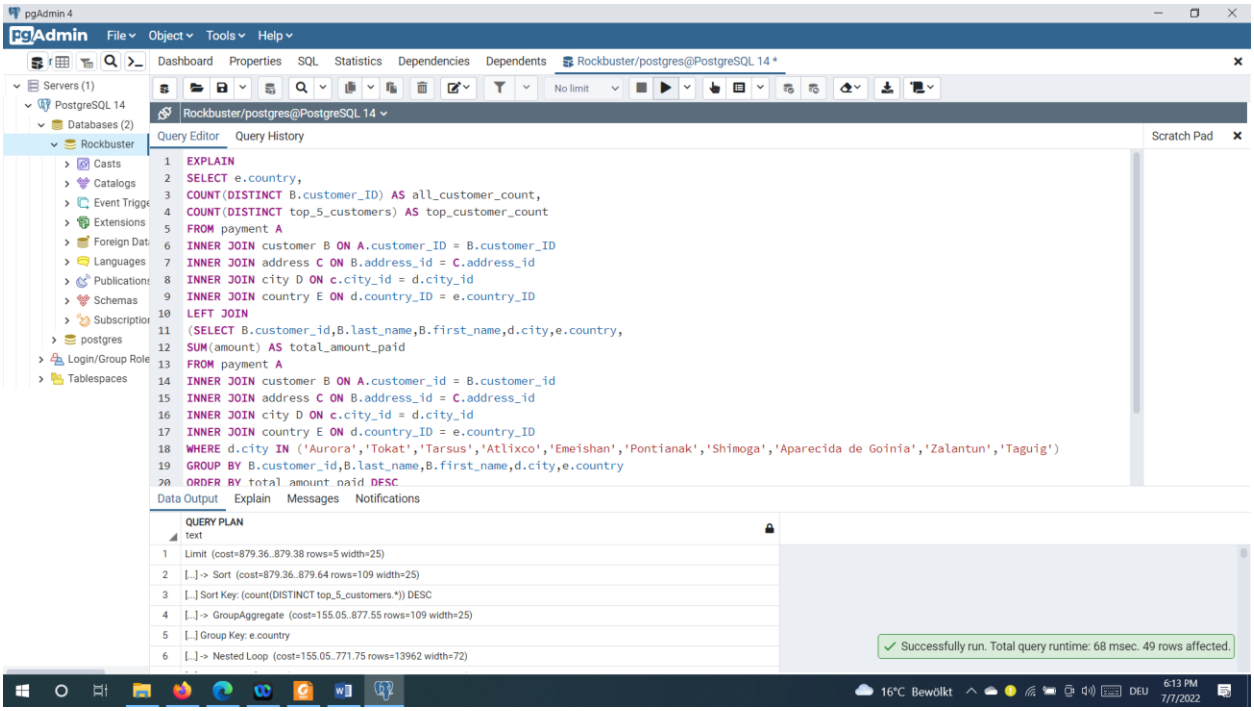
   I did not expect which approach would perform better.

2. Compare the costs of all the queries by creating query plans for each one.
3. 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.

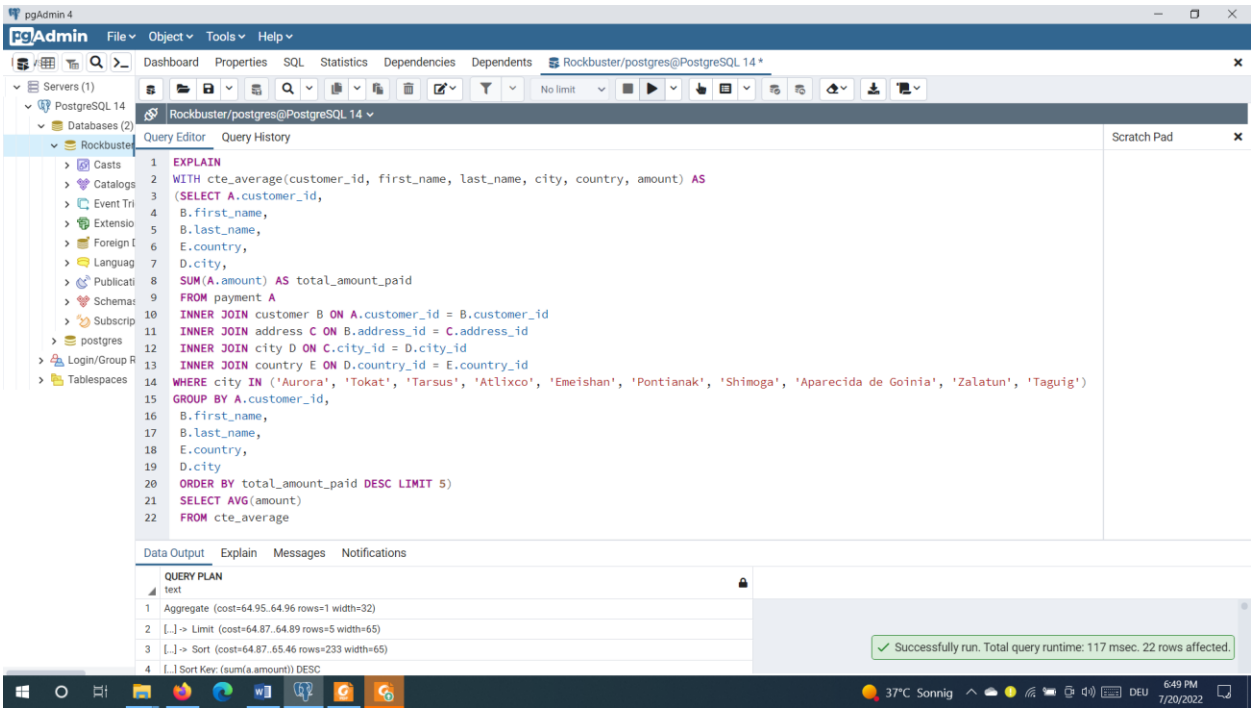
3.8.1



3.8.2


pgAdmin 4
File
Object
Tools
Help
Dashboard
Properties
SQL
Statistics
Dependencies
Dependents
Rockbuster/postgres@PostgreSQL 14 *
Servers (1)
PostgreSQL 14
Databases (2)
Rockbuster
Casts
Catalogs
Event Trigge
Extensions
Foreign Dat
Languages
Publications
Schemas
Subscriptio
postgres
Login/Group Role
Tablespaces
No limit
Rockbuster/postgres@PostgreSQL 14
Query Editor
Query History
Scratch Pad
EXPLAIN
SELECT e.country,
COUNT(DISTINCT B.customer_ID) AS all_customer_count,
COUNT(DISTINCT top_5_customers) AS top_customer_count
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
LEFT JOIN
(SELECT B.customer_id,B.last_name,B.first_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 d.city IN ('Aurora','Tokat','Tarsus','Atlixco','Emeishan','Pontianak','Shimoga','Aparecida de Goinia','Zalantun','Taguig')
GROUP BY B.customer_id,B.last_name,B.first_name,d.city,e.country
ORDER BY total_amount_paid DESC
Data Output
Explain
Messages
Notifications
QUERY PLAN
text
Limit (cost=879.36..879.38 rows=5 width=25)
[...] -> Sort (cost=879.36..879.64 rows=109 width=25)
[...] Sort Key: (count(DISTINCT top_5_customers.*)) DESC
[...] -> GroupAggregate (cost=155.05..877.55 rows=109 width=25)
[...] Group Key: e.country
[...] -> Nested Loop (cost=155.05..771.75 rows=13962 width=72)
Successfully run. Total query runtime: 68 msec. 49 rows affected.
16°C Bewölkt
DEU
6:13 PM
7/7/2022

3.9.1


pgAdmin 4
File Object Tools Help
Dashboard Properties SQL Statistics Dependencies Dependents Rockbuster/postgres@PostgreSQL 14 *
Rockbuster/postgres@PostgreSQL 14
Query Editor Query History
Scratch Pad
EXPLAIN
WITH cte_average(customer_id, first_name, last_name, city, country, amount) AS
(SELECT A.customer_id,
B.first_name,
B.last_name,
E.country,
D.city,
SUM(A.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', 'Tokat', 'Tarsus', 'Atlixco', 'Emeishan', 'Pontianak', 'Shimoga', 'Aparecida de Goinia', 'Zalatun', 'Taguig')
GROUP BY A.customer_id,
B.first_name,
B.last_name,
E.country,
D.city
ORDER BY total_amount_paid DESC LIMIT 5)
SELECT AVG(amount)
FROM cte_average
Data Output Explain Messages Notifications
QUERY PLAN
text
Aggregate (cost=64.95..64.96 rows=1 width=32)
[...] -> Limit (cost=64.87..64.89 rows=5 width=65)
[...] -> Sort (cost=64.87..65.46 rows=233 width=65)
[...] Sort Key: (sum(a.amount)) DESC
Successfully run. Total query runtime: 117 msec. 22 rows affected.

3.9.2

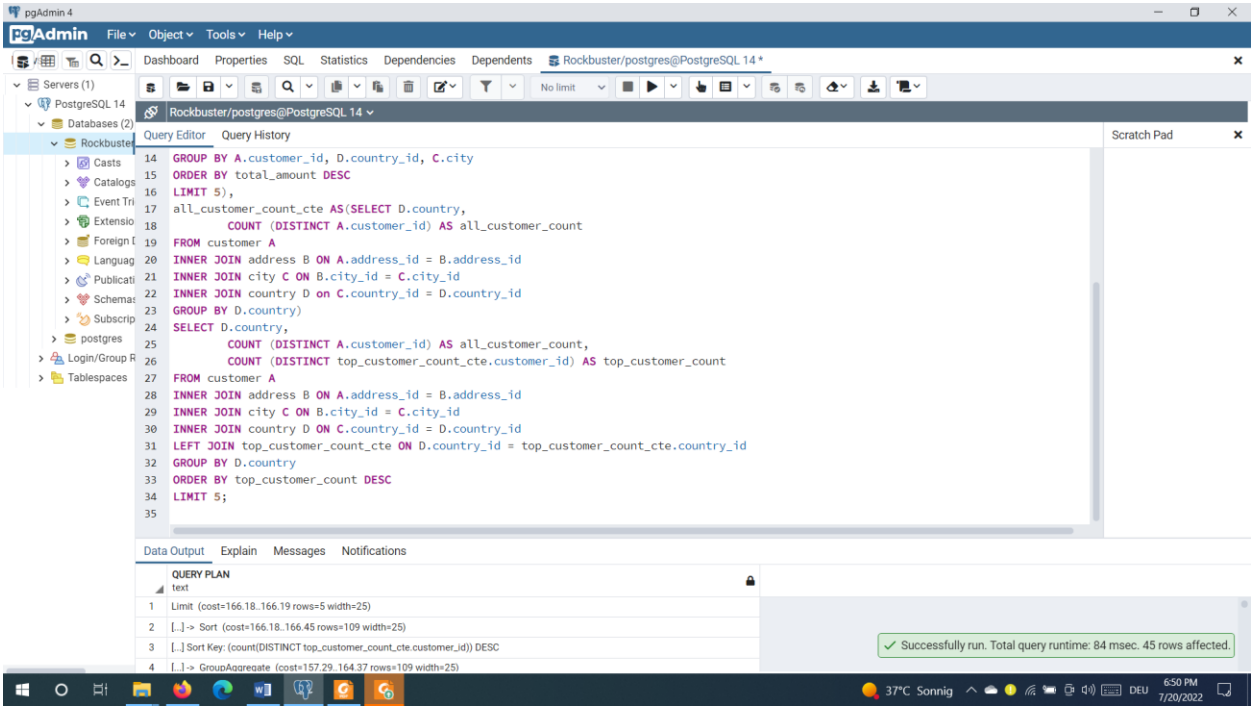

pgAdmin 4
File Object Tools Help
Dashboard Properties SQL Statistics Dependencies Dependents Rockbuster/postgres@PostgreSQL 14 *
Rockbuster/postgres@PostgreSQL 14
Query Editor Query History
Scratch Pad
GROUP BY A.customer_id, D.country_id, C.city
ORDER BY total_amount DESC
LIMIT 5),
all_customer_count_cte AS(SELECT D.country,
COUNT (DISTINCT A.customer_id) AS all_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_id = top_customer_count_cte.country_id
GROUP BY D.country
ORDER BY top_customer_count DESC
LIMIT 5;
Data Output Explain Messages Notifications
QUERY PLAN
text
Limit (cost=166.18..166.19 rows=5 width=25)
[...] -> Sort (cost=166.18..166.45 rows=109 width=25)
[...] Sort Key: (count(DISTINCT top_customer_count_cte.customer_id)) DESC
[...] -> GroupAggregate (cost=157.29..164.37 rows=109 width=25)
Successfully run. Total query runtime: 84 msec. 45 rows affected.

Table 1 contains the data about cost and speed of its queries.

Table 1.

QUERY	COST	SPEED (msec)
3.8.1	20284	78
3.8.2	879	68
3.9.1	64.95	117
3.9.2	166.18	84

4. Did the results surprise you? Write a few sentences to explain your answer.
   I am little surprised that COST of the Query 3.8.1 since it is significantly higher in compare with another queries.

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

For me writing the queries with CTE was harder.