

Katie Goyal

3.9 Common Table Expressions

Step 1:

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

Query	Query History
1	SELECT AVG(total_spent) AS average_paid_by_top_5_customers
2	FROM (
3	SELECT
4	A.customer_id,
5	B.first_name,
6	B.last_name,
7	E.country,
8	D.city,
9	SUM(A.amount) AS total_spent,
10	B.email
11	FROM payment A
12	INNER JOIN customer B ON A.customer_id = B.customer_id
13	INNER JOIN address C ON B.address_id = C.address_id
14	INNER JOIN city D ON C.city_id = D.city_id
15	INNER JOIN country E ON D.country_id = E.country_id
16	WHERE city IN ('Aurora', 'Atlixco', 'Xintai', 'Adoni', 'Dhule (Dhulla)', 'Khurasaki', 'Pingxiang', 'Sivas', 'Celaya', 'So Le
17	GROUP BY
18	A.customer_id,
19	B.first_name,
20	B.last_name,
21	A.amount,
22	B.email

Data Output	Messages	Notifications
average_paid_by_top_5_customers numeric		
1		34.930000000000000000

2. Copy-paste your CTEs and their outputs into your answers document

Query	Query History
1	WITH CustomerSpending AS (
2	SELECT
3	A.customer_id,
4	B.first_name,
5	B.last_name,
6	E.country,
7	D.city,
8	B.email,
9	SUM(A.amount) AS total_spent
10	FROM payment A
11	INNER JOIN customer B ON A.customer_id = B.customer_id
12	INNER JOIN address C ON B.address_id = C.address_id
13	INNER JOIN city D ON C.city_id = D.city_id
14	INNER JOIN country E ON D.country_id = E.country_id
15	WHERE D.city IN ('Aurora', 'Atlixco', 'Xintai', 'Adoni', 'Dhule (Dhulla)', 'Khurasaki', 'Pingxiang', 'Sivas', 'Celaya', 'So
16	GROUP BY
17	A.customer_id,
18	B.first_name,
19	B.last_name,
20	B.email,
21	E.country,
22	D.city

Data Output	Messages	Notifications
average_paid_by_top_5_customers numeric		
1		107.354000000000000000

3. I first broke down the original query into its component parts to better understand how it was handling data. I then reorganized these parts into a Common Table Expression (CTE) to make the query cleaner and easier to read. This restructuring not only made the query more logical but also improved its performance. After making these changes, I ran the revised query again to make sure it worked correctly and delivered the expected results efficiently.

Step 2:

I believe that using Common Table Expressions instead of subqueries will enhance performance, as they offer clearer readability. CTEs help to organize the structure of the query, potentially reducing its complexity and improving execution efficiency.

- Query 1 Task 3.8
 - Cost 34.90
 - Execution .030 SEC
- Query 1 Task 3.9
 - Cost 34.9
 - Execution .056 SEC
- Query 2 Task 3.8
 - Cost 270.33
 - Execution .075 SEC
- Query 2 task 3.9
 - Cost 107.35
 - Execution .069

I am not surprised by the results, as we observed that the queries using CTEs incurred higher cost and took longer to run.

Step 3:

Switching from subqueries to Common Table Expressions (CTEs) involved significant challenges, primarily in maintaining the integrity and complexity of the original query logic. Adapting to CTEs often required extensive restructuring of query flows and a deep understanding of the data, making the process time-consuming. Additionally, while CTEs enhance query readability and maintenance, they don't always improve performance; in fact, in large datasets, CTEs can lead to slower execution times. This necessitated rigorous performance testing and adjustments, including indexing and query hints, to ensure the system's efficiency was not compromised.