JANUARY 7, 2026

3.8 PERFORMING SUBQUERIES

JANELLA VALENCIA

Table of Contents

Step 1: Find the average amount paid by the top 5 customers

```
SELECT AVG(total_amount_paid) AS average_amount_paid
FROM
(SELECT B.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 D.city IN ('Aurora', 'Atlixco', 'Xintai', 'Adoni', 'Dhule (Dhulia)', 'Kurashiki',
'Pingxiang', 'Sivas', 'Celaya', 'So Leopoldo')
    AND E.country IN ('India', 'China', 'United States', 'Japan', 'Mexico', 'Brazil',
'Russian Federation', 'Philippines', 'Turkey', 'Indonesia')
GROUP BY B.customer_id, B.first_name, B.last_name, E.country, D.city
ORDER BY total_amount_paid DESC
LIMIT 5) AS average;
```

Output Step 1

	average_amount_paid numeric
1	107.3540000000000000

Step 2: Find out how many of the top 5 customers you identified in step 1 are based within each country

```
SELECT cnt1.country,
COUNT(DISTINCT cust1.customer_id) AS all_customer_count,
COUNT(DISTINCT top_5_customer.customer_id) AS top_customer_count
FROM customer AS cust1
INNER JOIN address AS addr1 ON cust1.address_id=addr1.address_id
INNER JOIN city AS cty1 ON addr1.city_id=cty1.city_id
INNER JOIN country AS cnt1 ON cty1.country_id=cnt1.country_id
LEFT JOIN (
    SELECT B.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 D.city IN ('Aurora', 'Atlixco', 'Xintai', 'Adoni', 'Dhule (Dhulia)', 'Kurashiki',
'Pingxiang', 'Sivas', 'Celaya', 'So Leopoldo')
    AND E.country IN ('India', 'China', 'United States', 'Japan', 'Mexico', 'Brazil',
'Russian Federation', 'Philippines', 'Turkey', 'Indonesia')
    GROUP BY B.customer_id, B.first_name, B.last_name, E.country, D.city
    ORDER BY total_amount_paid DESC
    LIMIT 5
) AS top_5_customer ON top_5_customer.country= cnt1.country
GROUP BY cnt1.country
ORDER BY all_customer_count DESC
LIMIT 5;
```

Output Step 2

	country character varying (50)	all_customer_count bigint	top_customer_count bigint
1	India	60	1
2	China	53	0
3	United States	36	1
4	Japan	31	0
5	Mexico	30	2

Step 3- Questions

Do you think step 1 and 2 could be done without using subqueries?

- For step 1, we're being asked to calculate the sum on payments for every customer and "top 5" customers by finding the average. To get to the answer, we need to use two aggregation steps that need the use of a subquery.
- For step 2, we're being asked to calculate 'total customer count per country' and at the same time filtering for the top 5 customers in specific cities. For this example I believe a subquery is needed to get your results because we're trying to preform two functions simultaneously.

When do you think subqueries are useful?

- I think subqueries are useful when the data is consistently changing. For example, wanting to know the average, or the top 5 customers/products/sales/ etc. Also, when you're being asked to do multiple aggregations in one query.