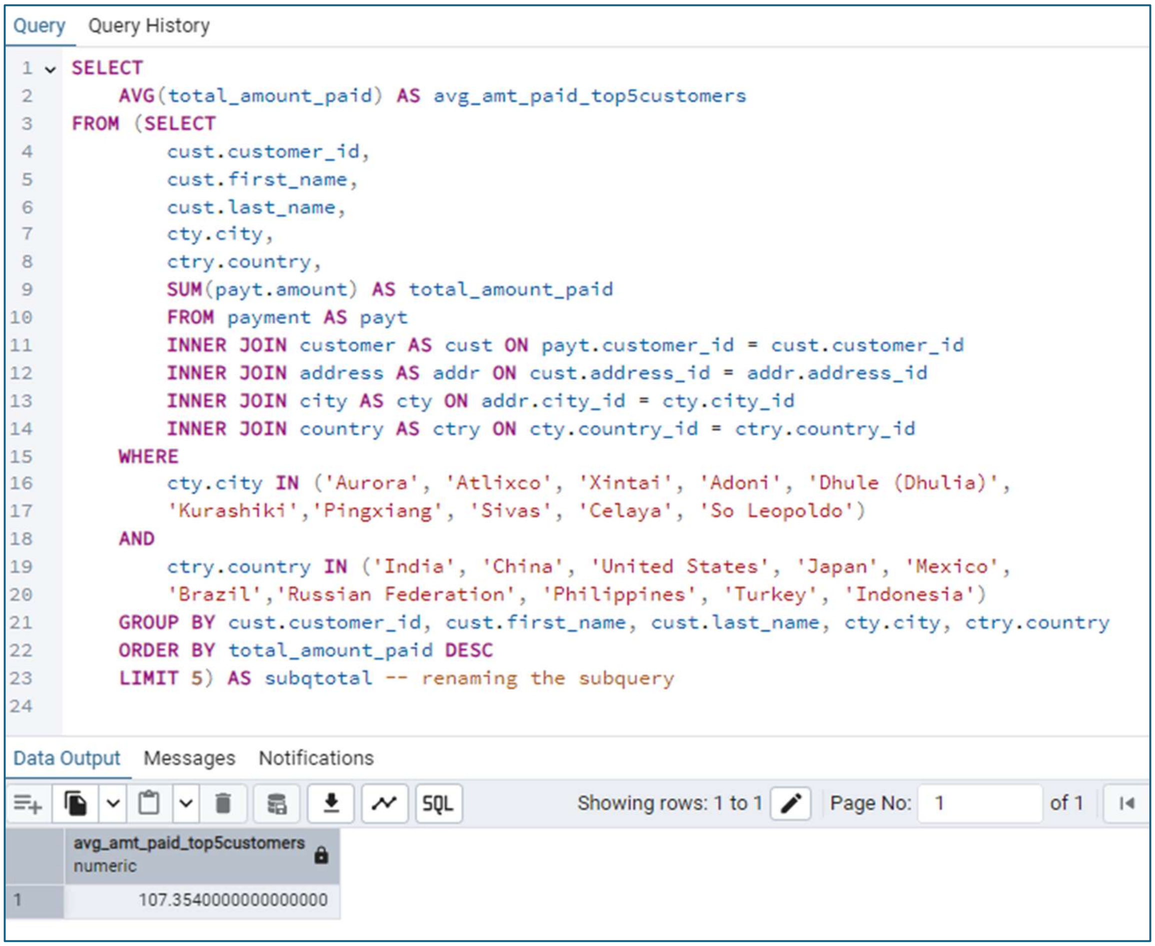
A3 - 3.8

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



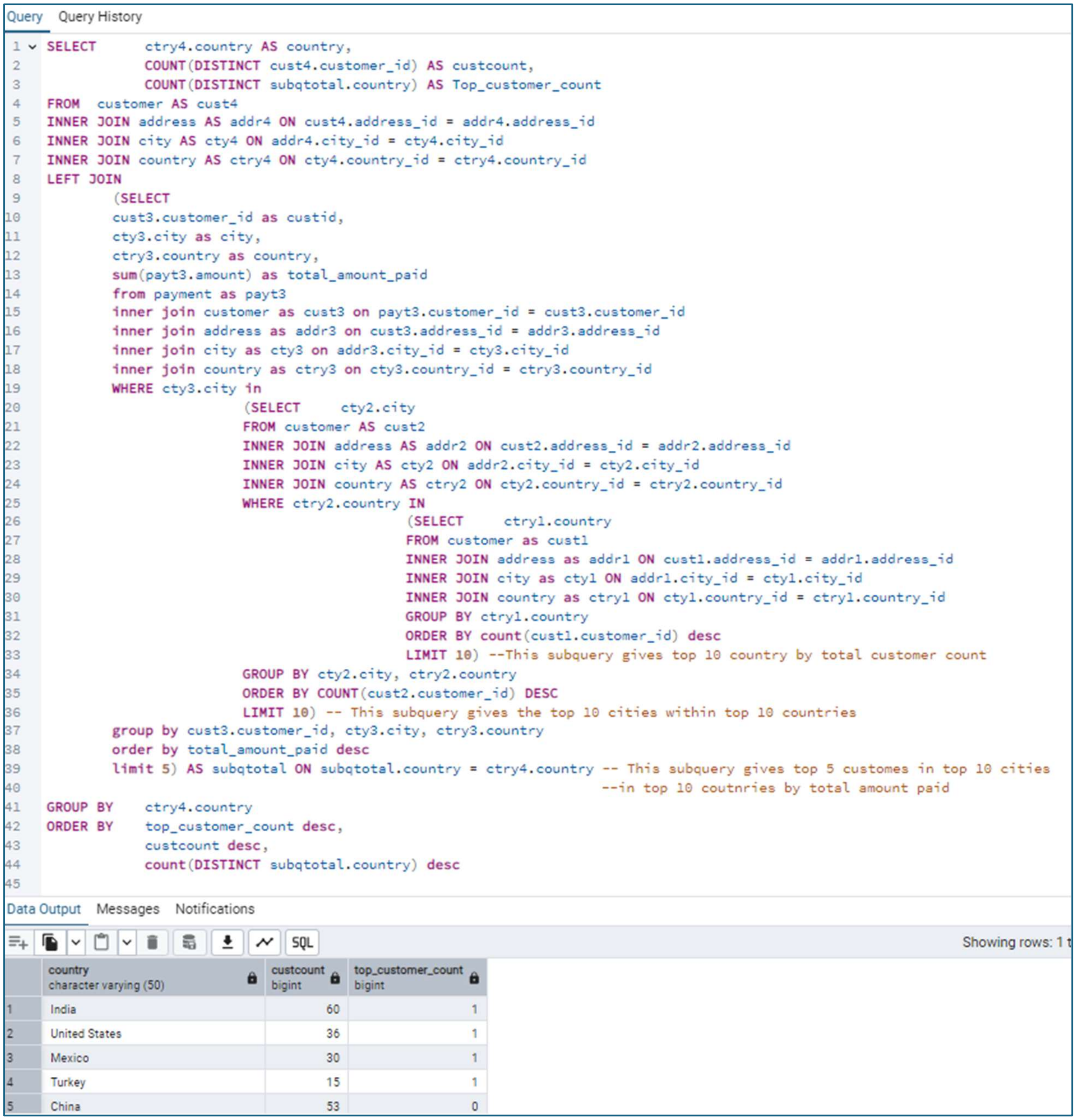
```sql
SELECT
    AVG(total_amount_paid) AS avg_amt_paid_top5customers
FROM (SELECT
        cust.customer_id,
        cust.first_name,
        cust.last_name,
        cty.city,
        ctry.country,
        SUM(payt.amount) AS total_amount_paid
        FROM payment AS payt
        INNER JOIN customer AS cust ON payt.customer_id = cust.customer_id
        INNER JOIN address AS addr ON cust.address_id = addr.address_id
        INNER JOIN city AS cty ON addr.city_id = cty.city_id
        INNER JOIN country AS ctry ON cty.country_id = ctry.country_id
    WHERE
        cty.city IN ('Aurora', 'Atlixco', 'Xintai', 'Adoni', 'Dhule (Dhulia)',
        'Kurashiki','Pingxiang', 'Sivas', 'Celaya', 'So Leopoldo')
    AND
        ctry.country IN ('India', 'China', 'United States', 'Japan', 'Mexico',
        'Brazil','Russian Federation', 'Philippines', 'Turkey', 'Indonesia')
    GROUP BY cust.customer_id, cust.first_name, cust.last_name, cty.city, ctry.country
    ORDER BY total_amount_paid DESC
    LIMIT 5) AS subqtotal -- renaming the subquery
```

	avg_amt_paid_top5customers 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.

Your final output should include 3 columns:

- “country”
- “all_customer_count” with the total number of customers in each country
- “top_customer_count” showing how many of the top 5 customers live in each country



```sql
SELECT      ctry4.country AS country,
            COUNT(DISTINCT cust4.customer_id) AS custcount,
            COUNT(DISTINCT subqtotal.country) AS Top_customer_count
FROM  customer AS cust4
INNER JOIN address AS addr4 ON cust4.address_id = addr4.address_id
INNER JOIN city AS cty4 ON addr4.city_id = cty4.city_id
INNER JOIN country AS ctry4 ON cty4.country_id = ctry4.country_id
LEFT JOIN
        (SELECT
        cust3.customer_id as custid,
        cty3.city as city,
        ctry3.country as country,
        sum(payt3.amount) as total_amount_paid
        from payment as payt3
        inner join customer as cust3 on payt3.customer_id = cust3.customer_id
        inner join address as addr3 on cust3.address_id = addr3.address_id
        inner join city as cty3 on addr3.city_id = cty3.city_id
        inner join country as ctry3 on cty3.country_id = ctry3.country_id
        WHERE cty3.city in
                        (SELECT     cty2.city
                        FROM customer AS cust2
                        INNER JOIN address AS addr2 ON cust2.address_id = addr2.address_id
                        INNER JOIN city AS cty2 ON addr2.city_id = cty2.city_id
                        INNER JOIN country AS ctry2 ON cty2.country_id = ctry2.country_id
                        WHERE ctry2.country IN
                                            (SELECT     ctry1.country
                                            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 ctry1 ON cty1.country_id = ctry1.country_id
                                            GROUP BY ctry1.country
                                            ORDER BY count(cust1.customer_id) desc
                                            LIMIT 10) --This subquery gives top 10 country by total customer count
                        GROUP BY cty2.city, ctry2.country
                        ORDER BY COUNT(cust2.customer_id) DESC
                        LIMIT 10) -- This subquery gives the top 10 cities within top 10 countries
        group by cust3.customer_id, cty3.city, ctry3.country
        order by total_amount_paid desc
        limit 5) AS subqtotal ON subqtotal.country = ctry4.country -- This subquery gives top 5 customes in top 10 cities
                                                                   --in top 10 coutnries by total amount paid
GROUP BY    ctry4.country
ORDER BY    top_customer_count desc,
            custcount desc,
            count(DISTINCT subqtotal.country) desc

```

	country character varying (50)	custcount bigint	top_customer_count bigint
1	India	60	1
2	United States	36	1
3	Mexico	30	1
4	Turkey	15	1
5	China	53	0

Step 3:

1. Write 1 to 2 short paragraphs on the following:

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

     Step 1 can be performed without subqueries as it is a calculation of average on one column. However, the ask in step 2 is relatively complicated.

     Step 2 requires various outputs to be established in a particular order before a final output can be extracted from that specific subset of the outputs. This can be either achieved by creating view or creating subqueries.

   - When do you think subqueries are useful?

     Subqueries are useful when a deep dive analysis is required to understand the finer details of any given data, especially while working on larger data set(s). Subqueries are also very effective when requirement dictates many aggregating functions.