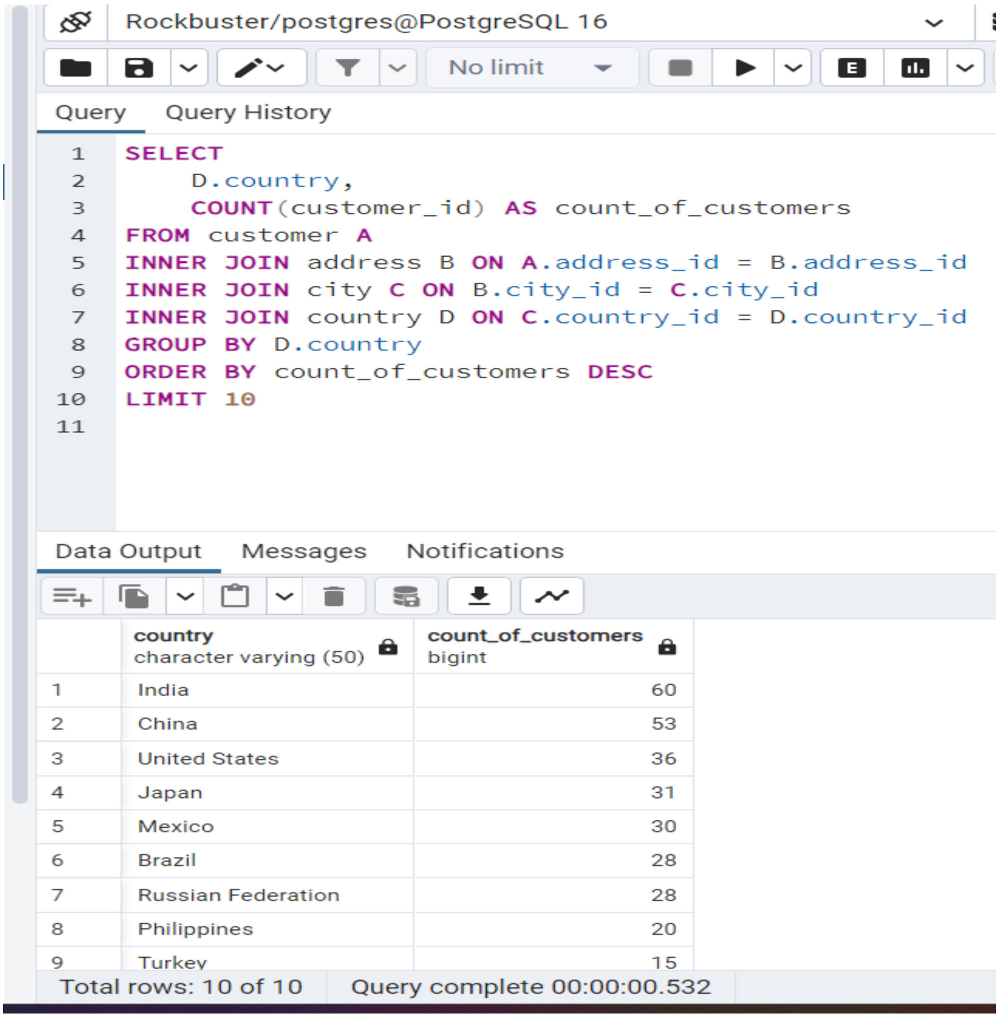
Olusola Olaniyi

Task 3.7 Joining Tables of Data

1A.
```
SELECT
   D.country,
   COUNT(customer_id) AS count_of_customers
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
ORDER BY count_of_customers DESC
LIMIT 10
```



```
SELECT
    D.country,
    COUNT(customer_id) AS count_of_customers
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
ORDER BY count_of_customers DESC
LIMIT 10

```

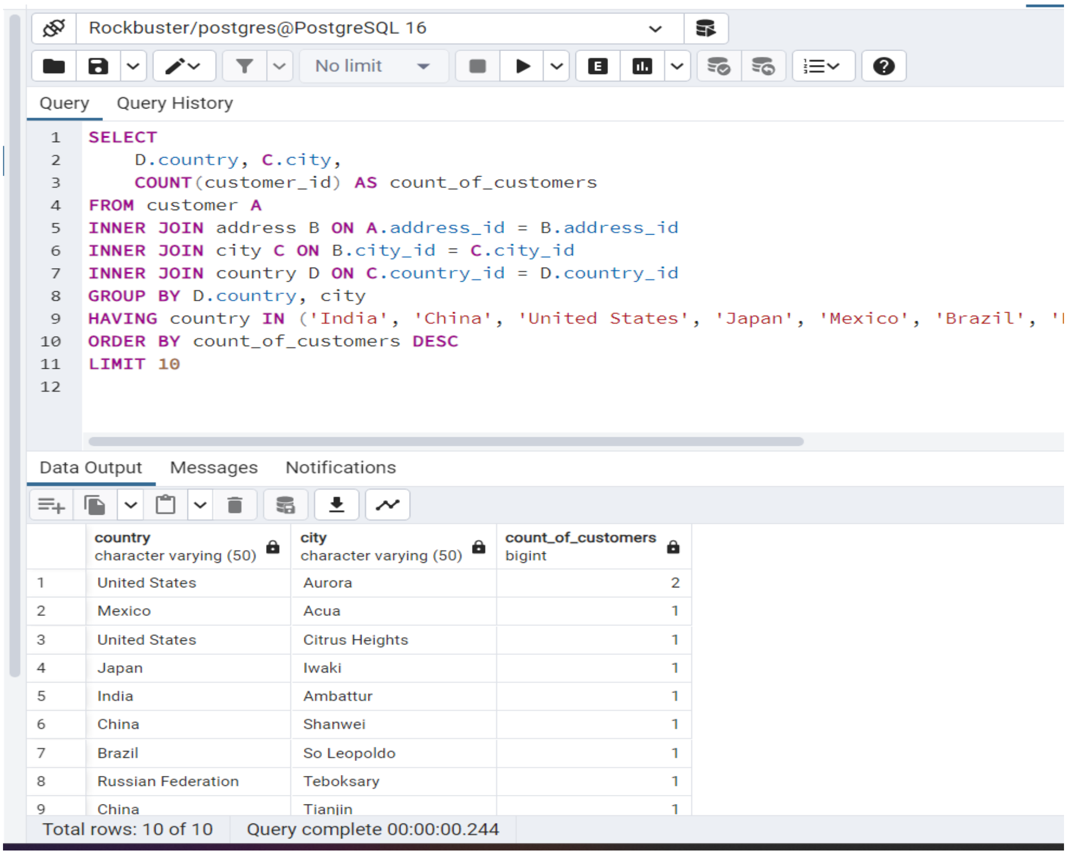
	country character varying (50)	count_of_customers bigint
1	India	60
2	China	53
3	United States	36
4	Japan	31
5	Mexico	30
6	Brazil	28
7	Russian Federation	28
8	Philippines	20
9	Turkey	15

Total rows: 10 of 10 Query complete 00:00:00.532

1B. The query uses JOIN statements to connect the "customer," "address," "city," and "country" tables based on their common keys (address_id, city_id, country_id). The COUNT(customer_id) is used to count the number of customers for each country. The GROUP BY D.country groups the results by country. The ORDER BY count_of_customers DESC orders the results in descending order based on the customer count. The LIMIT 10 ensures that only the top 10 countries are returned.

I will approach this query by identifying common keys for joining. As it is very crucial to understand the relationships between the tables. I will use the GROUP BY clause to aggregate the counts at the country level. I will use ORDER BY with LIMIT to retrieve the top 10 countries based on customer count.

2A. SELECT
  D.country, C.city,
  COUNT(customer_id) AS count_of_customers
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, city
HAVING country IN (‘India’, ‘China’, ‘United States’, ‘Japan’, ‘Mexico’, ‘Brazil’, ‘Russian Federation’, ‘Philippines’, ‘Turkey’, ‘Indonesia’)
ORDER BY count_of_customers DESC
LIMIT 10



```
SELECT
    D.country, C.city,
    COUNT(customer_id) AS count_of_customers
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, city
HAVING country IN ('India', 'China', 'United States', 'Japan', 'Mexico', 'Brazil', '
ORDER BY count_of_customers DESC
LIMIT 10
```

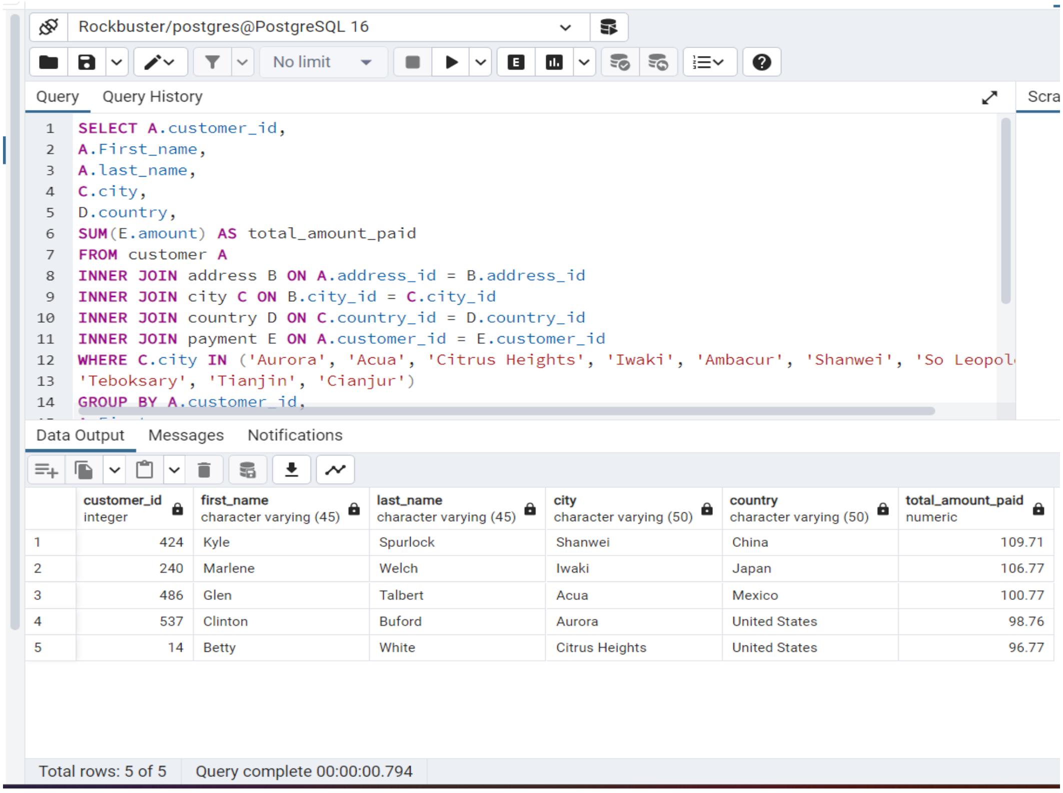
	country character varying (50)	city character varying (50)	count_of_customers bigint
1	United States	Aurora	2
2	Mexico	Acua	1
3	United States	Citrus Heights	1
4	Japan	Iwaki	1
5	India	Ambattur	1
6	China	Shanwei	1
7	Brazil	So Leopoldo	1
8	Russian Federation	Teboksary	1
9	China	Tianjin	1

Total rows: 10 of 10   Query complete 00:00:00.244

2B. I used the same query in step one and added C.city I used the INNER JOIN to include the city table. I grouped by city and country to get the customer count for each city within the top 10 countries.
Finally, the results are ordered by customer count in descending order, and the LIMIT 10 ensures only the top 10 cities are returned.
This approach ensures that the top 10 cities are identified within the context of the previously identified top 10 countries.

3A. SELECT A.customer_id,

```
A.First_name,
A.last_name,
C.city,
D.country,
SUM(E.amount) AS total_amount_paid
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
INNER JOIN payment E ON A.customer_id = E.customer_id
WHERE C.city IN ('Aurora', 'Acua', 'Citrus Heights', 'Iwaki', 'Ambacur', 'Shanwei', 'So
Leopoldo',
'Teboksary', 'Tianjin', 'Cianjur')
GROUP BY A.customer_id,
A.First_name,
A.last_name,
C.city,
D.country
ORDER BY total_amount_paid DESC
LIMIT 5
```



```
SELECT A.customer_id,
A.First_name,
A.last_name,
C.city,
D.country,
SUM(E.amount) AS total_amount_paid
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
INNER JOIN payment E ON A.customer_id = E.customer_id
WHERE C.city IN ('Aurora', 'Acua', 'Citrus Heights', 'Iwaki', 'Ambacur', 'Shanwei', 'So Leopol
'Teboksary', 'Tianjin', 'Cianjur')
GROUP BY A.customer_id,
```

	customer_id integer	first_name character varying (45)	last_name character varying (45)	city character varying (50)	country character varying (50)	total_amount_paid numeric
1	424	Kyle	Spurlock	Shanwei	China	109.71
2	240	Marlene	Welch	Iwaki	Japan	106.77
3	486	Glen	Talbert	Acua	Mexico	100.77
4	537	Clinton	Buford	Aurora	United States	98.76
5	14	Betty	White	Citrus Heights	United States	96.77

Total rows: 5 of 5 Query complete 00:00:00.794

3B.

I kept using the same INNER JOIN from step one. I added to SELECT A.customer_id because we are looking for top 5 customers. We need the information of the top 5 customers like First name, last name, city, and country. I used SUM as total amount paid (sum the total amount paid by top 5 customers).The result is grouped by customer and city to calculate the total amount paid by each customer and LIMIT 5 to ensure only the top 5 customers are returned.