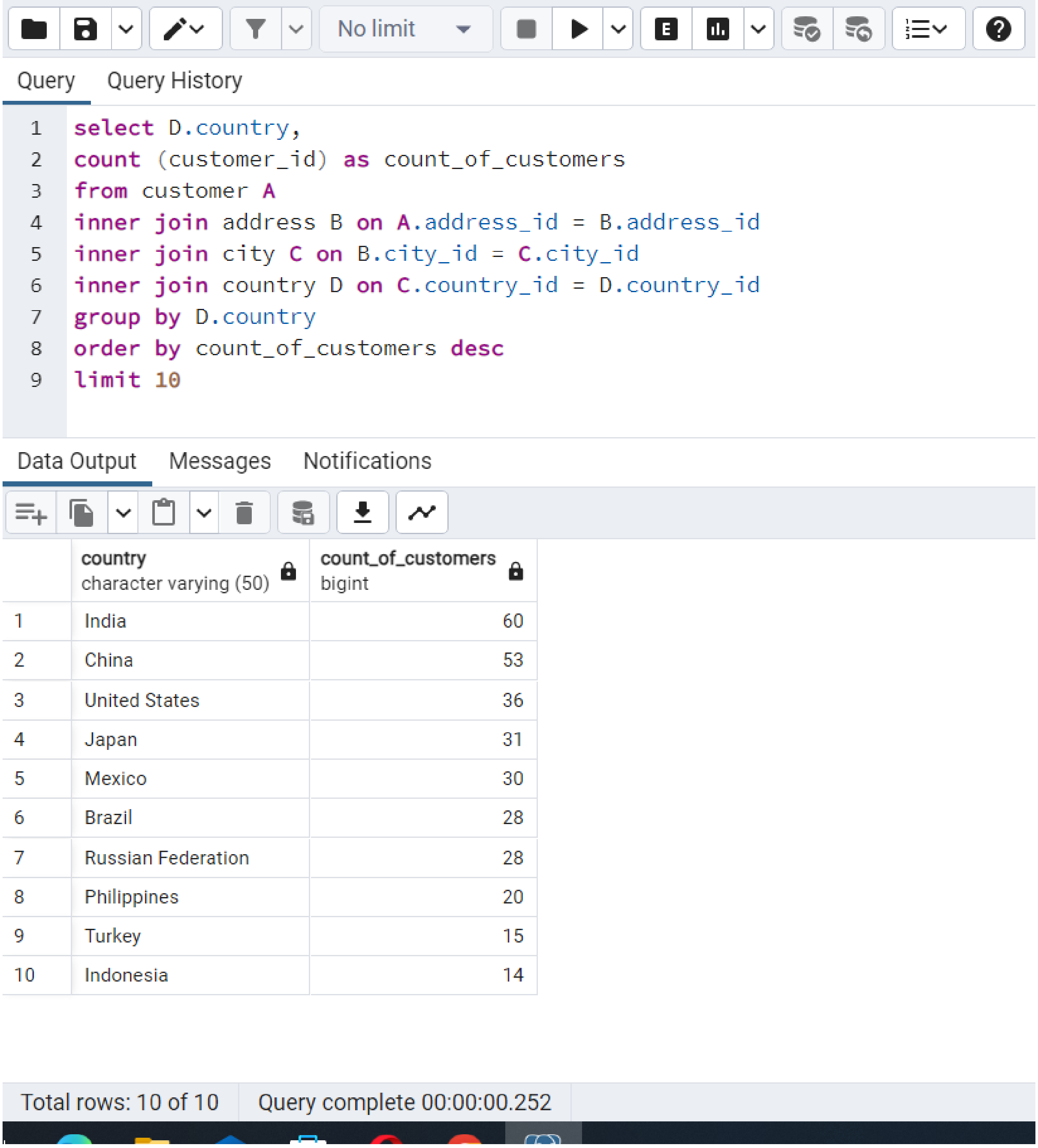
Surekha

Ach.3 Task 3.7

Exercise 3.7

Step 1:



```sql
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
10	Indonesia	14

Total rows: 10 of 10   Query complete 00:00:00.252

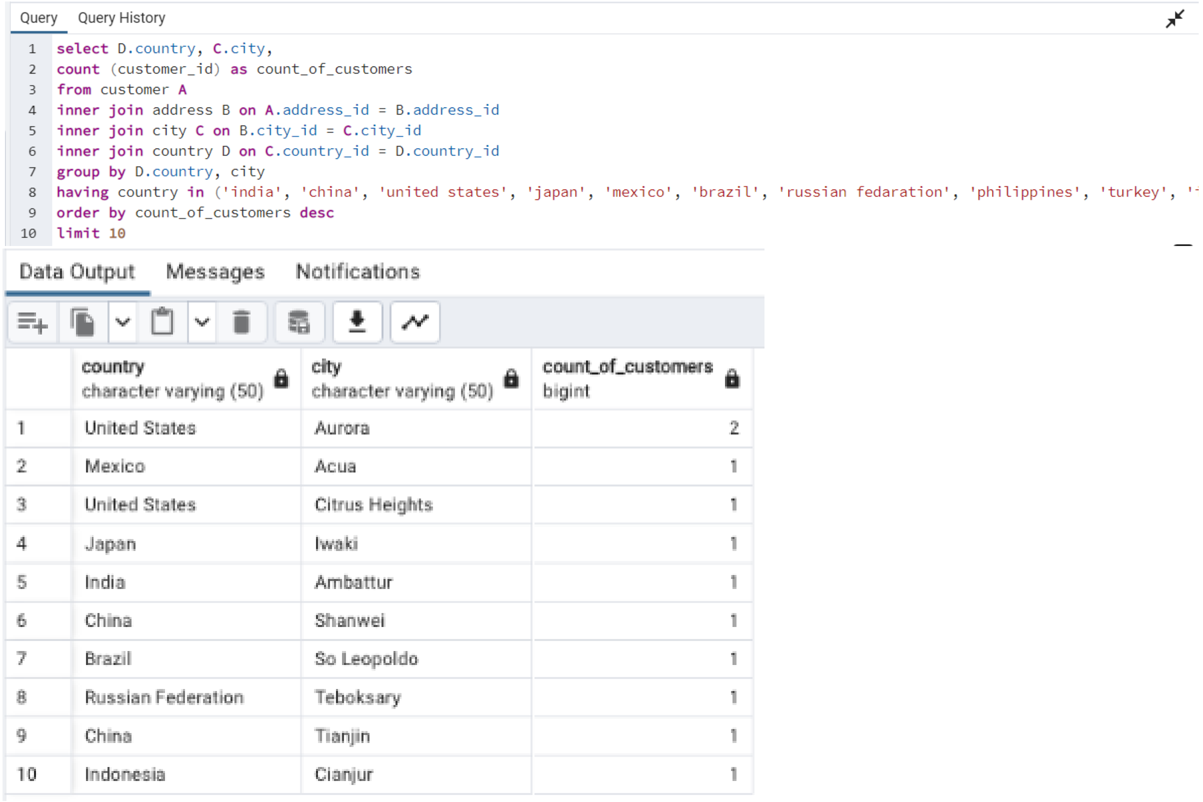
- For this question, we needed information from the customer table and from the country table. These tables do not have a direct link with one another given that they do not have a column in common. Because of this, we have to join other tables in order to be able to join customer and country together.
- For this query, we had to join the customer, address, city, and country tables together. The customer table had a link to the address table, the address table had a link to the city table, and the city table had a link to the country table.

- Since we need the information in the country column, we have selected that first in the query. Then, to find the top 10 countries based on customer numbers, we needed a count of the customer_id column, which we gave an alias to in order to improve readability.

- Now, to start the process of combining the country and customer tables, we had to combine the customer and address tables, then the city and address tables, and finally the country and city tables. I chose to use INNER JOIN because it is the most efficient JOIN command to use given that we only needed select information from each table.

- After joining all 4 tables to finally link the country and customers table together, we needed to group the results by country, and sort them in descending order so we can view the top 10 countries. To view only the top 10, the LIMIT 10 statement was used.

Step 2:



```sql
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 fedaration', 'philippines', 'turkey', '
order by count_of_customers desc
limit 10
```

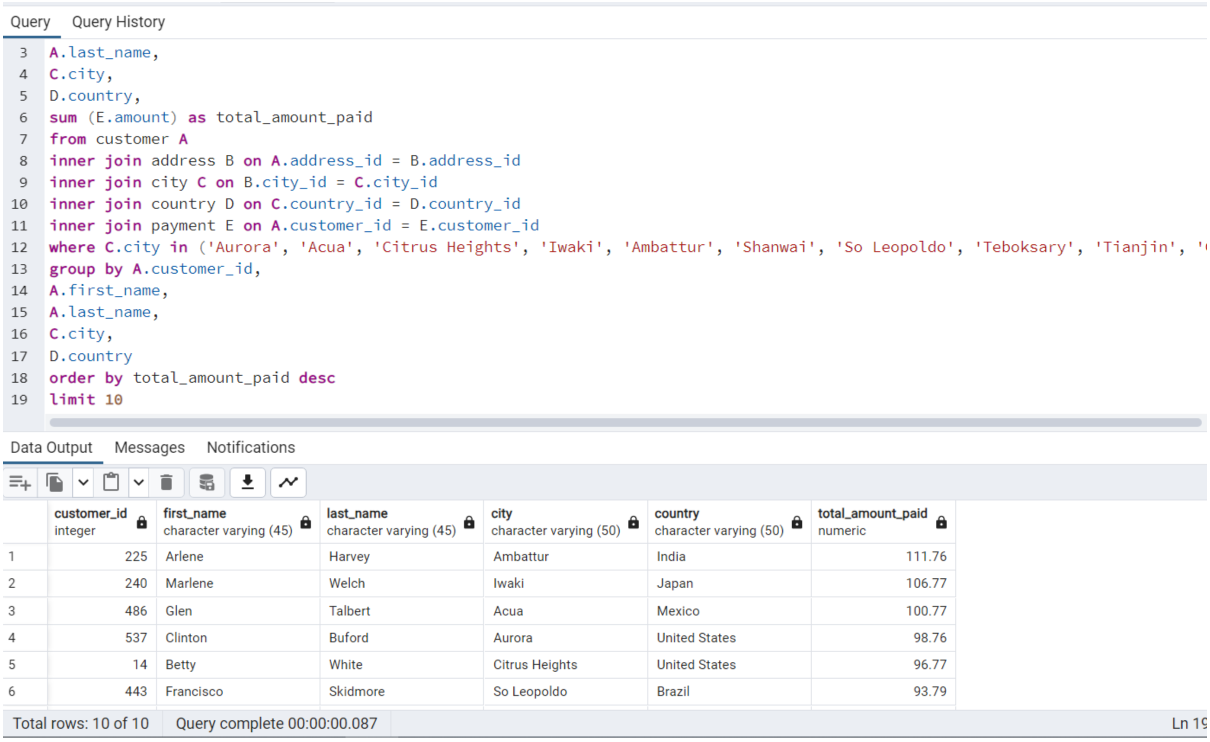
	country character varying (50)	city character varying (50)	count_of_customers bigint
1	United States	Aurora	2
2	Mexico	Acua	1
3	United States	Citrus Heights	1
4	Japan	Iwaki	1
5	India	Ambattur	1
6	China	Shanwei	1
7	Brazil	So Leopoldo	1
8	Russian Federation	Teboksary	1
9	China	Tianjin	1
10	Indonesia	Cianjur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

- I utilized the same query with some minor tweaks. Since we wanted to identify the top 10 cities within these countries, we needed to first select the city column to retrieve its data.

- Then, we need to specify that we are specifically looking at the top 10 countries in our previous query, so the statement does not look at other countries that are not top 10 in the table. To specify this, I used a WHERE clause to filter the data.

- Lastly, we want to group the results by city, in addition to the country because we want to identify the top 10 cities.

Step 3:



```
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', 'Ambattur', 'Shanwai', 'So Leopoldo', 'Teboksary', 'Tianjin', '
group by A.customer_id,
A.first_name,
A.last_name,
C.city,
D.country
order by total_amount_paid desc
limit 10
```

	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	225	Arlene	Harvey	Ambattur	India	111.76
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
6	443	Francisco	Skidmore	So Leopoldo	Brazil	93.79

Total rows: 10 of 10 Query complete 00:00:00.087 Ln 19

- I tweaked the previous query. The question asked for certain columns to be included in the output, so I selected them in the first section of the query. Each column is preceded by its table letter.

- We need all the amounts added up for the customer to get the total amount paid so the SUM command was used. I added a line to the query to join the new table “payment” with the customer table since they have the customer_id column in common.

- Since we are looking at the top 10 cities, the WHERE clause was changed to list the top 10 cities instead of the top 10 countries. I grouped the results by the necessary columns needed and ordered it by the total amount paid with a LIMIT 5 to retrieve the top 5 customers by total amount paid.