

Task 7 Join tables

Answers 3.7

Contents

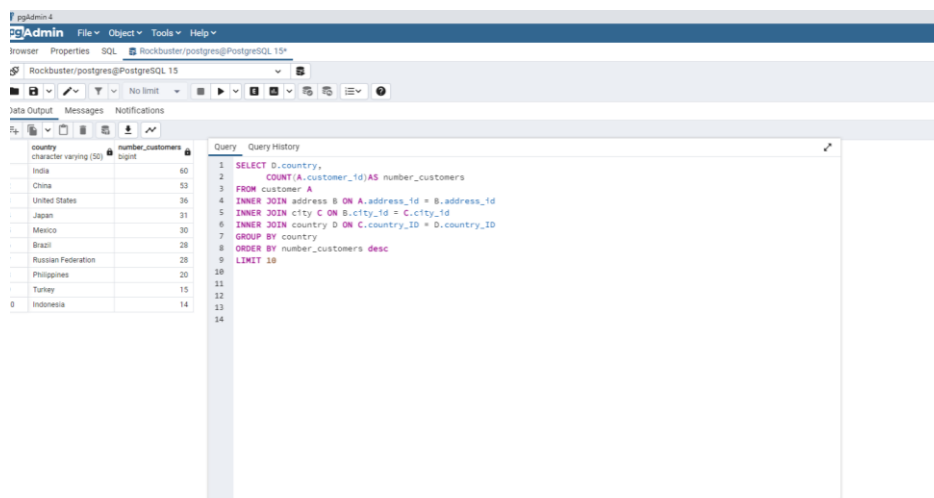
Task 7 Join tables	1
Answers 3.7.....	1
1. Write a query to find the top 10 countries for Rockbuster in terms of customer numbers.....	1
2. Write a query to find the top 10 cities within the top 10 countries identified in step 1. 2	
3. Write a query to find the top 5 customers in the top 10 cities who have paid the highest total amounts to Rockbuster.	3

1. Write a query to find the top 10 countries for Rockbuster in terms of customer numbers.

- Copy-paste your query and its output into your answers document.

Query:

```
SELECT D.country,  
       COUNT(A.customer_id)AS number_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 country  
ORDER BY number_customers desc  
LIMIT 10
```



The screenshot shows the pgAdmin 4 interface. The 'Query History' pane on the right displays the SQL query used to find the top 10 countries by customer count. The 'Data Output' pane on the left shows the results of the query as a table with two columns: 'country' and 'number_customers'.

country	number_customers
India	60
China	53
United States	36
Japan	31
Mexico	30
Brazil	28
Russian Federation	28
Philippines	20
Turkey	15
Indonesia	14

- Write a few sentences on how you approached this query and why.

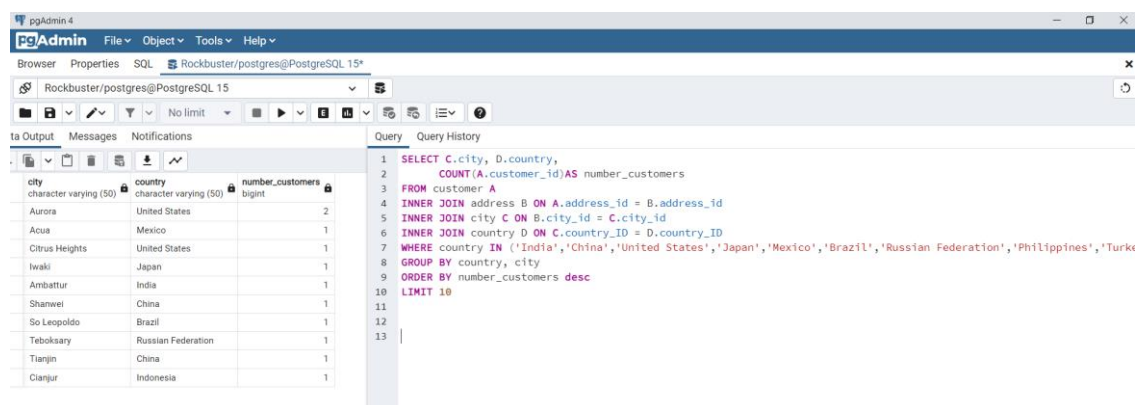
To approach this query, I first looked into the ERD to see which tables had the information required. I find that the customer table has the data on the customers, it has also a foreign key to the address table (i.e., address_id) which is linked to the city table, which is linked to the country table, (through the city_id and country_id respectively). So, I did a multiple JOIN query to joining data from these four tables. The JOIN statement used was the INNER join, because I aim to learn how which countries have the highest number of customers and every customer has an id which has an address and country, so I don't need another type of JOIN statement. To have the data grouped by countries I used GROUP BY. To order the results in relation to the countries with more customers, I used the ORDER BY statement to order the results of the COUNT (number of customer) in descending order. To obtain the top 10 countries with more customers, I used the LIMIT statement.

2. Write a query to find the top 10 cities within the top 10 countries identified in step 1.

- Copy-paste your query and its output into your answers document.

Query:

```
SELECT C.city, D.country,
       COUNT(A.customer_id)AS number_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
WHERE country IN ('India','China','United States','Japan','Mexico','Brazil','Russian
Federation','Philippines','Turkey','Indonesia')
GROUP BY country, city
ORDER BY number_customers desc
LIMIT 10
```



The screenshot shows the pgAdmin 4 interface. The 'Query' tab is active, displaying the SQL query. The 'Query History' tab shows the query was executed. The 'Table Output' tab displays the results of the query in a table format.

city	country	number_customers
Aurora	United States	2
Aqua	Mexico	1
Citrus Heights	United States	1
Iwaki	Japan	1
Ambattur	India	1
Shanwei	China	1
So Leopoldo	Brazil	1
Teboksary	Russian Federation	1
Tianjin	China	1
Cianjur	Indonesia	1

- Write a short explanation of how you approached this query and why.

To approach this query, I used a similar query than the previous. However, instead of selecting as just the country, I also selected the city. I did a multiple JOIN query that enables joining data from four tables, to deliver the top 10 cities in terms of number of

customers and used WHERE with the operator IN to identify the countries from which the query should retrieve the cities.

To have the data grouped by city and country, I used GROUP BY, and the ORDER BY of the number of customers in descending order, which gives me the top 10 cities with more customers, (also using the LIMIT statement).

3. Write a query to find the top 5 customers in the top 10 cities who have paid the highest total amounts to Rockbuster.

- Tip: After the join syntax, you'll need to use the WHERE clause with an operator, followed by GROUP BY and ORDER BY. Your output should include the following columns: Customer ID, Customer First Name and Last Name, Country, City, Total Amount Paid.
- Copy-paste your query and its output into your answers document.

Query:

```
SELECT B.customer_id,  
       B.first_name,  
       B.last_name,  
       D.city,  
       E.country,  
       SUM (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 city IN ('Aurora','Acua','Citrus Heights','Iwaki','Ambattur','Shanwei','So  
Leopoldo','Teboksary','Tianjin','Cianjur')  
GROUP BY B. customer_id, B.first_name, B.last_name, D.city, E.country  
ORDER BY total_amount_paid DESC  
LIMIT 5
```

The screenshot shows the DBeaver Admin interface with a query executed on a PostgreSQL database. The query results are displayed in a table with 6 columns: customer_id, first_name, last_name, city, country, and total_amount_paid. The results show the top 5 customers based on the total amount paid, filtered by specific cities. The SQL query is visible in the right-hand pane.

	customer_id	first_name	last_name	city	country	total_amount_paid
1	225	Arlene	Harvey	Ambattur	India	111.76
2	424	Kyle	Spurlock	Shanwei	China	109.71
3	240	Marlene	Welch	Iwaki	Japan	106.77
4	486	Glen	Talbert	Acua	Mexico	100.77
5	537	Clinton	Buford	Aurora	United States	98.76

```
1 SELECT B.customer_id,  
2       B.first_name,  
3       B.last_name,  
4       D.city,  
5       E.country,  
6       SUM (amount)AS total_amount_paid  
7 FROM payment A  
8 INNER JOIN customer B ON A.customer_id=B.customer_id  
9 INNER JOIN address C ON B.address_id=C.address_id  
10 INNER JOIN city D ON C.city_id=D.city_id  
11 INNER JOIN country E ON D.country_id=E.country_id  
12 WHERE city IN ('Aurora','Acua','Citrus Heights','Iwaki','Ambattur','Shanwei','So Leopoldo','Teboksary','Tianjin','Cianjur')  
13 GROUP BY B. customer_id, B.first_name, B.last_name, D.city, E.country  
14 ORDER BY total_amount_paid DESC  
15 LIMIT 5  
16  
17  
18  
19
```