

Answer 3.7: Joining Tables of Data

1. Write a query to find the top 10 countries for Rockbuster in terms of customer numbers. (Tip: you'll have to use **GROUP BY** and **ORDER BY**, both of which follow the join.)

1a. Copy-paste your query and its output into your answers document.

Query

Query History

```
1 SELECT D.country,
2 COUNT(customer_id)AS Count_of_customer_numbers
3 FROM Customer A
4 INNER JOIN address B ON A. address_id=B. address_id
5 INNER JOIN city C ON B. city_id= C.city_id
6 INNER JOIN country D ON C.country_id= D.country_id
7 GROUP BY country
8 ORDER BY COUNT(A.customer_id )DESC
9 LIMIT 10;
```

Data output

Messages

Notifications

	country character varying (50)	count_of_customer_numbers 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 00:00:00.089

1b. Write a few sentences on how you approached this query and why. It's important that you can explain your thought process when writing queries, especially for future interviews.

- ❖ I used the multiple joins and inner joins to connect the multiple tables. Inner joins were used because I only needed information from two tables but used the multiple joins approach because the two tables where the information was are not connected.

Firstly, use the SELECT command to find information from the Country table and counted the customer ID to know how many customers for each country before using the join syntax to connect four tables to get the information needed from the two tables (customer and country).

I group it by country since we wanted to know the top 10 countries where Rockbuster has the most customers. Then, I order the table by the count of the customers' ids in descending order to be able to see the top country with the most customer. Finally, I limit it to the top 10 according to Rockbuster's request.

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

2a. Copy-paste your query and its output into your answers document.

The screenshot shows a SQL query editor with a toolbar at the top containing icons for file operations, filters, and execution. Below the toolbar, there are tabs for 'Query' and 'Query History'. The 'Query' tab is active, displaying a SQL query. Below the query, there are tabs for 'Data output', 'Messages', and 'Notifications'. The 'Data output' tab is active, showing a table with 10 rows of results. The table has three columns: 'city', 'country', and 'count_of_customer_numbers'. The data is as follows:

	city character varying (50)	country character varying (50)	count_of_customer_numbers bigint
1	Aurora	United States	2
2	Bhusawal	India	1
3	Shivapuri	India	1
4	Cianjur	Indonesia	1
5	Kuwana	Japan	1
6	Acua	Mexico	1
7	Saint Louis	United States	1
8	So Leopoldo	Brazil	1
9	Iwaki	Japan	1
10	Eskisehir	Turkey	1

At the bottom of the interface, there is a status bar showing 'Total rows: 10 of 10' and 'Query complete 00:00:00.093'.

2b. Write a short explanation of how you approached this query and why.

- ❖ I approached this task 2 as task 1, the difference is that I selected city as well so the table can show city name, country name and count of customer ids. To display the 10 top cities within the top 10 countries, I included the WHERE and IN commands to specify in the query where(country) and in which countries (top 10 countries) the top 10 cities should be picked from. I also group by both city and country.

3. Write a query to find the top 5 customers in the top 10 cities who have paid the highest total amounts to Rockbuster. The customer team would like to reward them for their loyalty!

- **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, and Total Amount Paid.

The screenshot shows a SQL query editor with a toolbar at the top. The query is as follows:

```
1 SELECT A.amount,
2 B.customer_id,
3 B.first_name,
4 B.last_name,
5 D.city,
6 E.country,
7 SUM(amount) AS Total_amount_paid
8 FROM payment A
9 INNER JOIN customer B ON A.customer_id=B.customer_id
10 INNER JOIN address C ON B.address_id= C.address_id
11 INNER JOIN city D ON C.city_id= D.city_id
12 INNER JOIN country E ON D.country_id= E.country_id
13 WHERE city IN ('Aurora','Bhusawal','Shivapuri', 'Cianjur',
14 'Kuwana','Acua','Saint Louis','So Leopoldo','Iwaki','Eskisehir' )
15 GROUP BY A.amount, B.customer_id, B.first_name, B.last_name, D.city, E.country
16 ORDER BY SUM(amount) DESC
17 LIMIT 5;
```

Below the query editor, there are tabs for "Data output", "Messages", and "Notifications". The "Data output" tab is active, showing a table with 8 columns: amount, customer_id, first_name, last_name, city, country, and total_amount_paid. The table contains 5 rows of data.

	amount numeric (5,2)	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	4.99	486	Glen	Talbert	Acua	Mexico	59.88
2	4.99	240	Marlene	Welch	Iwaki	Japan	39.92
3	4.99	537	Clinton	Buford	Aurora	United States	39.92
4	4.99	202	Carla	Gutierrez	Bhusawal	India	34.93
5	4.99	507	Edgar	Rhoads	Eskisehir	Turkey	34.93

At the bottom of the interface, it says "Total rows: 5 of 5" and "Query complete 00:00:00.097".