

ANSWERS 3.7

- 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.)

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

The screenshot shows the pgAdmin 4 interface. On the left, the 'customer' table is selected under the 'country' schema. The main pane displays the following SQL query:

```

1 SELECT D.country,
2       COUNT(customer_id) as total_number_of_customers
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(customer_id) DESC
9 LIMIT 10

```

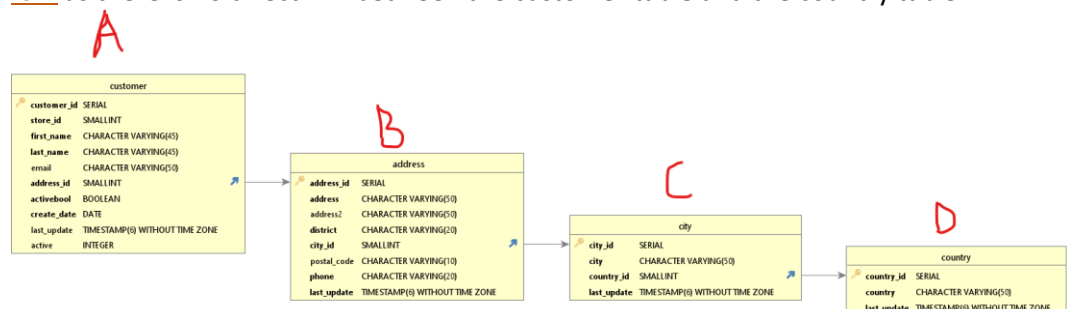
The 'Data Output' pane shows the results of the query:

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

Total rows: 10 of 10. Query complete 00:00:00.314. Ln 9, Col 9.

❖ 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.

For executing this query, I first used DB visualizer to find out which tables need to be used for the answer, by way of ERD, this gives us the relevant tables we need. As we see below that we only need Customer count and Country so we need to use Multiple Join as there is no direct link between the customer table and the country table.



Thereafter, I used the Inner Join (assigning tables A to D) to write the query as the data we need is limited. I grouped the data by Country and ORDER BY count in DESC to get the number of customers from top to bottom and thereafter setting LIMIT to 10 as we needed data only of the top 10 countries.

ANSWERS 3.7

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.

The screenshot shows the pgAdmin 4 interface. On the left, the 'Tables (17)' folder is expanded, showing the 'customer' table. The 'Columns (10)' for 'customer' are listed: customer_id, store_id, first_name, last_name, email, address_id, activebool, create_date, last_update, and active. The 'Queries' folder is also expanded, showing a query named 'Rockbuster/postgres@PostgreSQL 15'. The query text is as follows:

```

1 SELECT D.country,
2       C.city,
3       COUNT(customer_id) as total_number_of_customers
4 FROM customer A
5 INNER JOIN address B ON A.address_id = B.address_id
6 INNER JOIN city C ON B.city_id = C.city_id
7 INNER JOIN country D ON C.country_id = D.country_id
8 WHERE country IN ('India','China','United States','Japan','Mexico','Brazil','Russian Federation','Philippines','Turkey','Indonesia')
9 GROUP BY city,country
10 ORDER BY count(customer_id) DESC
11 LIMIT 10

```

The 'Data Output' tab shows the results of the query. The table has three columns: 'country', 'city', and 'total_number_of_customers'. The results are as follows:

country	city	total_number_of_customers
United States	Aurora	2
Mexico	Atlixco	1
China	Xintai	1
India	Adoni	1
India	Dhule (Dhulia)	1
Japan	Kurashiki	1
China	Pingxiang	1
Turkey	Sivas	1
Mexico	Celaya	1
Brazil	So Leopoldo	1

The status bar at the bottom indicates 'Total rows: 10 of 10' and 'Query complete 00:00:00.645'.

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

For finding out the corresponding cities for the Top 10 countries we found earlier, I selected the city table (C), and specifically asked for the relevant countries by way of the WHERE clause with IN operator for including the relevant country names.

ANSWERS 3.7

- 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!

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

Taking the payment table as A, Customer as B, Address as C, City as D and Country as E, we have executed the below query. The ERD helps us in identifying the tables where the relevant data lies.

The screenshot shows the pgAdmin 4 interface with a SQL query executed in the 'Query' tab. The query is as follows:

```

2  B.first_name,
3  B.last_name,
4  D.city,
5  E.country,
6  SUM(A.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','Atlixco','Xintai','Adoni','Dhule(Dhulia)','Kurashiki','Pingxiang','Sivas','Celaya','So Leopoldo')
13 GROUP BY B.customer_id,
14          B.first_name,
15          B.last_name,
16          D.city,
17          E.country
18 ORDER BY Total_Amount_Paid DESC
19 LIMIT 5

```

The 'Data Output' tab shows the results of the query:

	customer_id	first_name	last_name	city	country	total_amount_paid
1	84	Sara	Perry	Atlixco	Mexico	128.7
2	518	Gabriel	Harder	Sivas	Turkey	108.75
3	587	Sergio	Stanfield	Celaya	Mexico	102.76
4	537	Clinton	Buford	Aurora	United States	98.76
5	367	Adam	Gooch	Adoni	India	97.8

Total rows: 5 of 5 Query complete 00:00:01.481

