

Answers 3.7

Rockbuster's management team would like to know the top 10 countries where Rockbuster customers are based so they can focus on building a better brand image in those markets. Follow the instructions below to find out how you can help!

Directions:

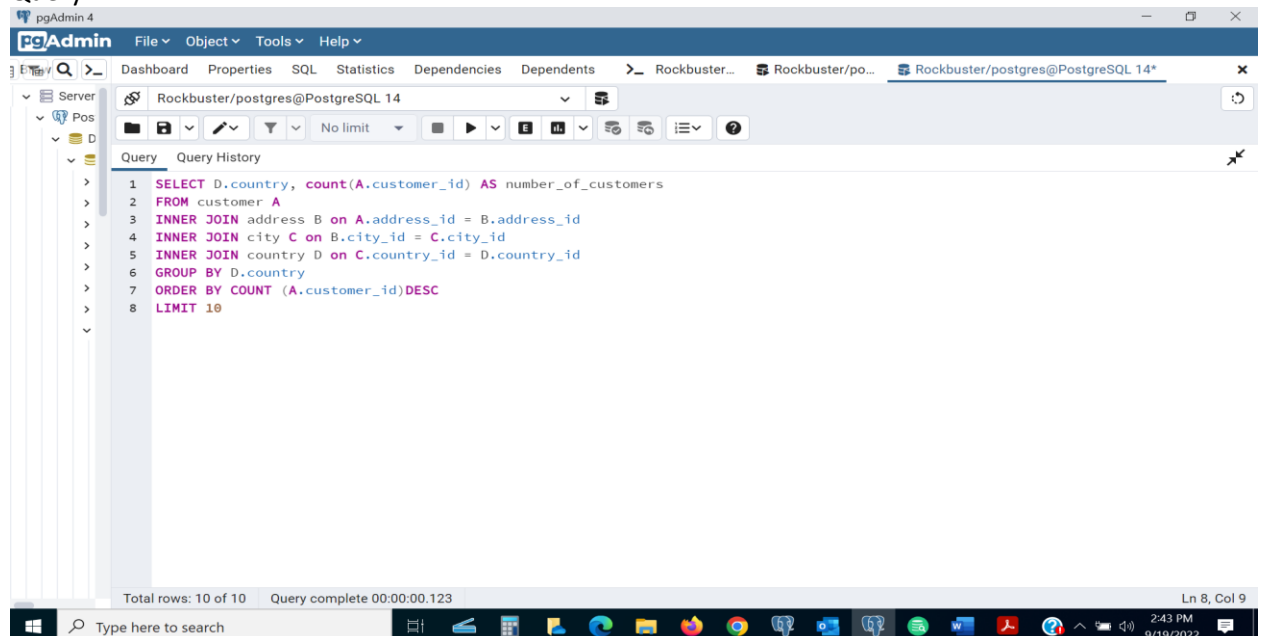
In this Task you'll get to practice everything you learned in the Exercise. Said simply, you'll need to write a couple of queries combined with joins between the tables address, country, city, customer and payment using their common keys. Create a new text document and call it "Answers 3.7." You'll save your queries, outputs and written answers in this document, as you've done in previous tasks.

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

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.

- o Looked at data dictionary / ERD that shows tables and relationships
- o Customer table and country table have data on country and customer ID.
- o Created query Select the country and counted number of customer ID's named column "number of customers"
- o Joined Tables with inner join which looks for matching values in the join column(namely the foreign key or primary key) – customer, address, city, and country tables
- o Group by to sum each country, order by count, descending order customer id to display largest number of customers (top 10)

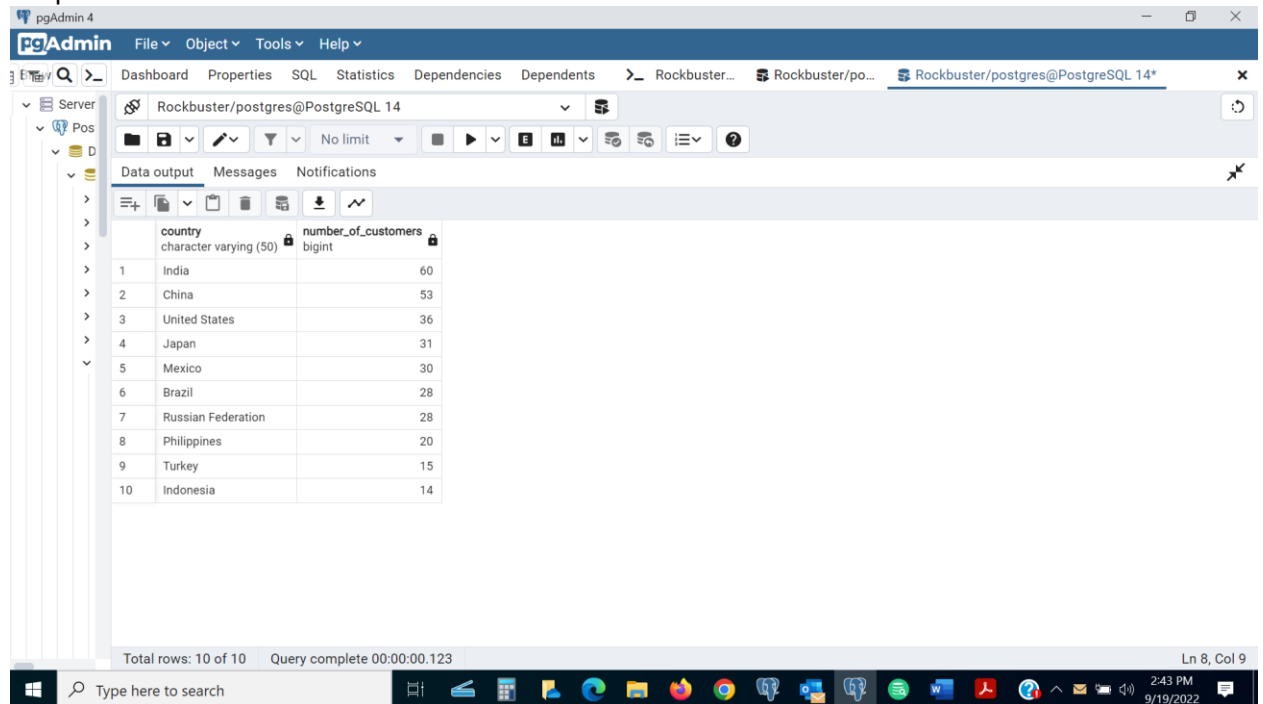
Query1:



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

Total rows: 10 of 10 Query complete 00:00:00.123 Ln 8, Col 9

Output1:



The screenshot shows the pgAdmin 4 interface. The 'Data output' tab is active, displaying a table with 10 rows. The table has two columns: 'country' (character varying (50)) and 'number_of_customers' (bigint). The data is as follows:

	country	number_of_customers
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

The status bar at the bottom indicates 'Total rows: 10 of 10' and 'Query complete 00:00:00.123'. The taskbar at the very bottom shows the Windows search bar and various application icons.

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.
 - Write a short explanation of how you approached this query and why.
 - Looked at data dictionary / ERD that shows tables and relationships
 - Modified step 1 Query to include city variable from City table, then used WHERE D.country IN ('India','China','United States','Japan','Mexico','Brazil','Russian Federation','Philippines','Turkey','Indonesia') **the top countries identified from first query above**
 - Then used GROUP BY C.city, D.country for display of records to show the top 10 cities within the top 10 countries that were identified in step 1

Query2:

The screenshot shows the pgAdmin 4 interface with a SQL query entered in the Query editor. The query is as follows:

```

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

```

The status bar at the bottom indicates "Total rows: 10 of 10" and "Query complete 00:00:00.221".

Output2:

The screenshot shows the pgAdmin 4 interface with the "Data output" tab selected. The query results are displayed in a table with the following columns: city, country, and number_of_customers.

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

The status bar at the bottom indicates "Total rows: 10 of 10" and "Query complete 00:00:00.221".

- 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, Total Amount Paid.
- Copy-paste your query and its output into your answers document.
- Looked at data dictionary / ERD that shows tables and relationships
- Modified step 2 Query to include customer id, last name, first name, city, country and amount AS amount paid in the select statement.
- Added the payment table to join on customer id
- Used WHERE D.city IN ('Aurora', 'Atlixco', 'Xintai', 'Adoni', 'Dhule (Dhulia)', 'Kurashiki', 'Pingxiang', 'Sivas', 'Celaya', 'So Leopoldo')**the top cities identified from 2nd query above**
- Then used GROUP BY A.customer_id, A.first_name, A.last_name, D.city, E.country for display of records to show the top 5 customers and amounts paid.

QUERY3:

The screenshot shows the pgAdmin 4 interface with a SQL query entered in the Query Editor. The query is as follows:

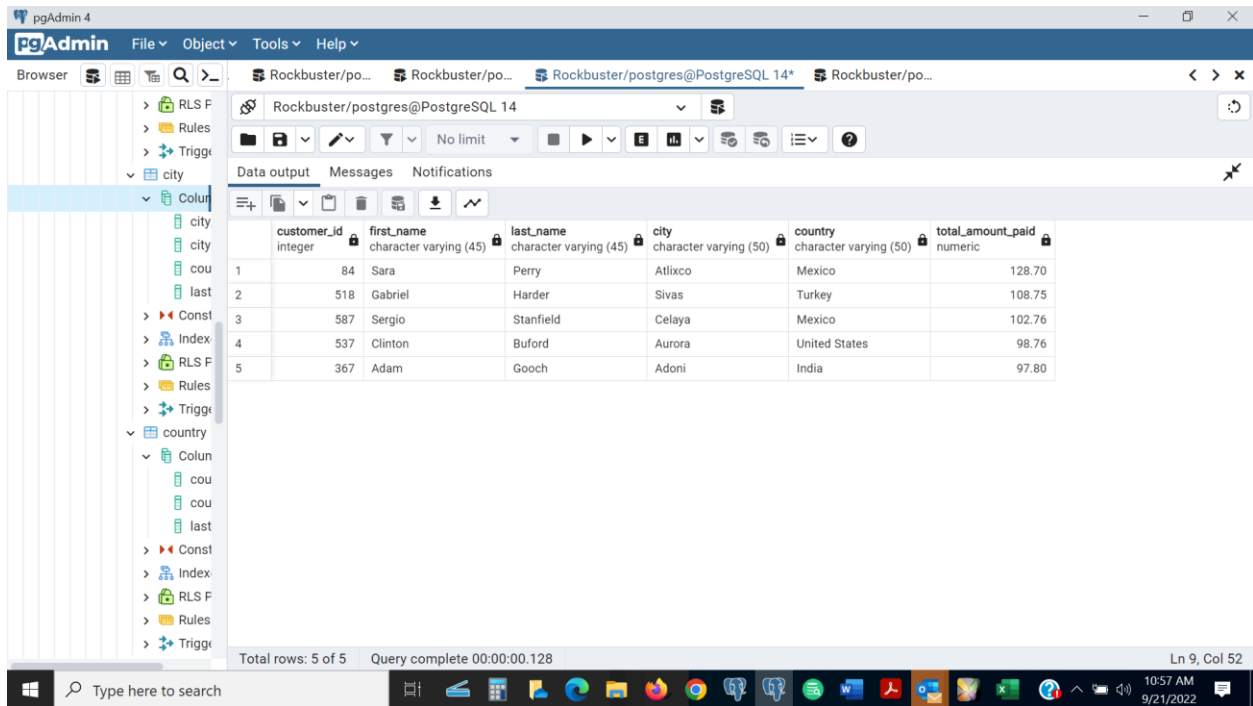
```

1  SELECT A.customer_id,
2      A.first_name,
3      A.last_name,
4      D.city,
5      E.country,
6      SUM(B.amount) AS total_amount_paid
7  FROM customer A
8  INNER JOIN payment B ON A.customer_id = B.customer_id
9  INNER JOIN address C ON A.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 D.city IN ('Aurora', 'Atlixco', 'Xintai', 'Adoni', 'Dhule (Dhulia)',
13                 'Kurashiki', 'Pingxiang', 'Sivas', 'Celaya', 'So Leopoldo')
14 GROUP BY A.customer_id, A.first_name, A.last_name, D.city, E.country
15 ORDER BY total_amount_paid DESC
16 LIMIT 5;

```

The interface also shows a tree view on the left with the database structure, and a status bar at the bottom indicating 'Total rows: 5 of 5' and 'Query complete 00:00:00.128'.

OUTPUT3:



The screenshot shows the pgAdmin 4 interface. The left sidebar displays a tree view of the database structure, including 'city' and 'country' tables. The main pane shows the 'Data output' tab for a query executed on the 'Rockbuster/postgres@PostgreSQL 14' database. The query result is a table with 5 rows and 7 columns: customer_id, first_name, last_name, city, country, and total_amount_paid. The status bar at the bottom indicates 'Total rows: 5 of 5' and 'Query complete 00:00:00.128'.

	customer_id	first_name	last_name	city	country	total_amount_paid
1	84	Sara	Perry	Atlitico	Mexico	128.70
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	97.80

4. Save your "Answers 3.7" document as a PDF and upload it here for your tutor to review.