Exercise 3.7 Joining Tables of Data

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

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

Query | Query History

```
SELECT C.country,
       COUNT(D.customer_id) AS customer_count
FROM customer D
JOIN address A ON A.address_id = D.address_id
JOIN city B ON A.city_id = B.city_id
JOIN country C ON C.country_id = B.country_id

GROUP BY C.country
ORDER BY COUNT(D.customer_id) DESC
LIMIT 10
```

	country character varying (50)	customer_count 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


```
SELECT C.country,
       COUNT(D.customer_id) AS customer_count
FROM customer D
JOIN address A ON A.address_id = D.address_id
JOIN city B ON A.city_id = B.city_id
JOIN country C ON C.country_id = B.country_id

GROUP BY C.country
ORDER BY COUNT(D.customer_id) DESC
```

LIMIT 10

- Write a few sentences on how you approached this query and why. You must be able to explain your thought process when writing queries, especially for future interviews.

Rockbuster wants to determine the number of customers in each country. To achieve this, I need to include both the country name and the total count of customers associated with each country in the records. To retrieve this information, I will link the relevant tables using their key fields. Then, I will group the data by country and sort the results in descending order based on the customer count. Finally, I will limit the output to the top 10 countries.

2. Next, write a query to identify the top 10 cities that fall within the top 10 countries you identified in step 1. (Hint: the top 10 cities can be in any of the countries identified—you don't need to create a separate list for each country.)

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

Query | Query History

```sql
SELECT C.country,
       B.city,
       COUNT(B.city_id) AS city_count,
       COUNT(D.customer_id) AS customer_count
FROM customer D

JOIN address A ON A.address_id = D.address_id
JOIN city B ON A.city_id = B.city_id
JOIN country C ON C.country_id = B.country_id

WHERE C.country IN (
    SELECT C.country
    FROM customer D
    JOIN address A ON A.address_id = D.address_id
    JOIN city B ON A.city_id = B.city_id
    JOIN country C ON C.country_id = B.country_id
    GROUP BY C.country
    ORDER BY COUNT(D.customer_id) DESC
    LIMIT 10
    )


GROUP BY C.country, B.city
ORDER BY COUNT (D.customer_id) DESC
LIMIT 10
```

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

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

I first need to filter out the top 10 countries, then extract and sort the top 10 cities from that result, all while ensuring the relevant tables are properly linked through their key fields.

Now write a query to find the top 5 customers from the top 10 cities who've 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.
- Copy-paste your query and its output into your answers document.

```
SELECT D.customer_id,
       D.first_name,
       D.last_name,
       B.city,
       C.country,
       D.email,
       SUM(E.amount) AS total_paid
FROM customer D
JOIN address A ON A.address_id = D.address_id
JOIN city B ON A.city_id = B.city_id
JOIN country C ON C.country_id = B.country_id
JOIN payment E ON E.customer_id = D.customer_id

WHERE B.city IN(

    SELECT B.city
    FROM customer D
    JOIN address A ON A.address_id = D.address_id
    JOIN city B ON A.city_id = B.city_id
    JOIN country C ON C.country_id = B.country_id

WHERE C.country IN (
    SELECT C.country
    FROM customer D
    JOIN address A ON A.address_id = D.address_id
    JOIN city B ON A.city_id = B.city_id
    JOIN country C ON C.country_id = B.country_id
    GROUP BY C.country
    ORDER BY COUNT(D.customer_id) DESC
    LIMIT 10
    )
GROUP BY C.country, B.city
ORDER BY COUNT (D.customer_id) DESC
```

LIMIT 10
)

GROUP BY B.city, C,country, D.last_name, D.first_name, D.email, D.customer_id
ORDER BY SUM(E.amount)DESC
LIMIT 5

Query | Query History

```sql
SELECT D.customer_id,
       D.first_name,
       D.last_name,
       B.city,
       C.country,
       D.email,
       SUM(E.amount) AS total_paid
FROM customer D
JOIN address A ON A.address_id = D.address_id
JOIN city B ON A.city_id = B.city_id
JOIN country C ON C.country_id = B.country_id
JOIN payment E ON E.customer_id = D.customer_id

WHERE B.city IN(

    SELECT B.city
    FROM customer D
    JOIN address A ON A.address_id = D.address_id
    JOIN city B ON A.city_id = B.city_id
    JOIN country C ON C.country_id = B.country_id

WHERE C.country IN (
    SELECT C.country
    FROM customer D
    JOIN address A ON A.address_id = D.address_id
    JOIN city B ON A.city_id = B.city_id
    JOIN country C ON C.country_id = B.country_id
    GROUP BY C.country
    ORDER BY COUNT(D.customer_id) DESC
```

Showing rows: 1 to 5 | Page No: 1

	customer_id integer	first_name character varying (45)	last_name character varying (45)	city character varying (50)	country character varying (50)	email character varying (50)	total_paid numeric
1	225	Arlene	Harvey	Ambattur	India	arlene.harvey@sakilacustomer.org	111.76
2	424	Kyle	Spurlock	Shanwei	China	kyle.spurlock@sakilacustomer.org	109.71
3	240	Marlene	Welch	Iwaki	Japan	marlene.welch@sakilacustomer....	106.77
4	486	Glen	Talbert	Acua	Mexico	glen.talbert@sakilacustomer.org	100.77
5	537	Clinton	Buford	Aurora	United States	clinton.buford@sakilacustomer.org	98.76