3.7: Joining Tables of Data

1. What tables are included in the database and how they're called:

```
SELECT * FROM pg_catalog.pg_tables
```

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. 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.)

- *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.*

In ERD I've chosen the necessary columns for the query:

address_id and customer_id from customer
address_id and city_id from address
city_id and country_id from city
country_id and country from country

To calculate the number of customers in each country I counted the number of customer id's and gave the column an alias ("number_of_customers"). Then I've used INNER JOIN to join customer data with country data, using address and city tables, as there is not a direct link between customer and country tables. GROUPed BY country, as we're interested in numbers in each country, ORDERed BY the number of customers in descending order, and LIMITed to 10 results, since we need top 10 countries.

```
SELECT D.country,
COUNT (A.customer_id) AS number_of_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 COUNT (A.customer_id) DESC
LIMIT 10
```

	country character varying (50)	number_of_customers 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

2. *Write a query to find the top 10 cities within the top 10 countries identified in step 1.*
    - *Write a short explanation of how you approached this query and why.*

I've added WHERE and IN syntax to my previous query to select only the cities we I've identified the answer to the 1st question. At the end I've changed GROUP BY clause from country to city:

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

	city character varying (50)	number_of_customers bigint
1	Aurora	2
2	Tokat	1
3	Tarsus	1
4	Atlixco	1
5	Emeishan	1
6	Pontianak	1
7	Shimoga	1
8	Aparecida de Goinia	1
9	Zalantun	1
10	Taguig	1

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, Total Amount Paid.*

I've revised ERD again and added payment data from the payment table to my query to extract the total amount paid.

customer_id, first_name, last_name, address_id from customer
customer_id and amount from payment
address_id and city_id from address
city_id, city, country_id from city
country_id and country from country

Similar to my previous query I've used city names in WHERE clause (instead of countries). LIMITed cusomers by 5.

```
SELECT A.customer_id, A.first_name, A.last_name, D.country,
C.city,
SUM (E.amount) AS total_amount_paid
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
INNER JOIN payment E ON A.customer_id=E.customer_id
WHERE C.city IN
('Aurora','Tokat','Tarsus','Atlixco','Emeishan','Pontianak','Shi
moga','Aparecida de Goinia','Zalantun','Taguig')
GROUP BY A.customer_id, A.first_name, A.last_name, C.city,
D.country
ORDER BY total_amount_paid DESC
LIMIT 5
```

	customer_id integer	first_name character varying (45)	last_name character varying (45)	country character varying (50)	city character va
1	566	Casey	Mena	Turkey	Tokat
2	84	Sara	Perry	Mexico	Atlixco
3	506	Leslie	Seward	Indonesia	Pontianak
4	389	Alan	Kahn	China	Emeishan
5	537	Clinton	Buford	United States	Aurora