Exercise 3.7 - Joining Tables of Data

Step 1 :

Write a query to find the top 10 countries for Rockbuster in terms of customer numbers.

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

Query | Query History

```
1  SELECT co.country AS customer_country,
2   COUNT(cu.customer_id) AS number_of_customer
3  FROM customer AS cu
4  INNER JOIN address AS ad ON cu.address_id = ad.address_id
5  INNER JOIN city AS ci ON ad.city_id = ci.city_id
6  INNER JOIN country AS co ON ci.country_id = co.country_id
7  GROUP BY country
8  ORDER BY number_of_customer DESC
9  LIMIT 10;
```

Data output | Messages | Notifications

	customer_country character varying (50)	number_of_customer bigint
1	India	60
2	China	53
3	United States	36
4	Japan	31
5	Mexico	30

- 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.

→

1. Review ERD

I started off by reviewing the ERD created in an earlier task (Task 3.2).

2. Identification of tables

As we need to know the top 10 countries in terms of the total number of customers, therefore I looked at the ERD to see from which table(s) can we get this information from - and in this case that would be the customer and country tables. However, in order to reach this step, there are some other tables that have some common keys that connect these tables which I'll refer to in the next step.

3. Query formation

Once I saw that there were multiple tables involved, I checked the common key in the remaining tables and decided on using INNER JOIN as the type of JOIN for the query. The reason to choose this type of JOIN is that this is a query where we need only limited information from two tables.
In order to write the query, I assigned the following tables to the corresponding alphabet to execute this query.
A = customer , B = address, C = city, D = country

4. Correct Sequence in the Query

Applying the correct sequence for the additional requirements of the following:

- GROUP BY
- ORDER BY with the function (DESC or ASC),
- LIMIT

Step 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.

Query | Query History

```
1  SELECT ci.city AS customer_city,
2   co.country AS customer_country,
3   COUNT(cu.customer_id) AS number_of_customer
4  FROM customer AS cu
5  INNER JOIN address AS ad ON cu.address_id = ad.address_id
6  INNER JOIN city AS ci ON ad.city_id = ci.city_id
7  INNER JOIN country AS co ON ci.country_id = co.country_id
8  WHERE country IN ('India', 'China', 'United States', 'Japan', 'Mexico',
9   'Brazil', 'Russian Federation', 'Philippines', 'Turkey', 'Indonesia')
10 GROUP BY city, country
11 ORDER BY number_of_customer DESC
12 LIMIT 10;
```

Data output | Messages | Notifications

	customer_city character varying (50)	customer_country character varying (50)	number_of_customer bigint
1	Aurora	United States	2
2	Atlixco	Mexico	1
3	Xintai	China	1
4	Adoni	India	1

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

- Keeping the steps in mind that I followed in Step 1, I added the city table to the SELECT part of the query as we want to know the top 10 cities within the top 10 countries.

- In the query, when using the WHERE clause, I restricted the country to the top 10 countries from Step 1 by using the IN operator.

- While trying to arrive at the result, initially I didn't use it and the countries were not the same as mentioned in Step 1 therefore the IN operator with the WHERE clause is necessary in order to reach at this result.

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

Query | Query History

```
 4   ci.city,
 5   co.country,
 6   SUM(pay.amount) AS total_amount_paid
 7  FROM customer as cu
 8  INNER JOIN address AS ad ON cu.address_id = ad.address_id
 9  INNER JOIN city AS ci ON ad.city_id = ci.city_id
10  INNER JOIN country AS co ON ci.country_id = co.country_id
11  INNER JOIN payment AS pay ON cu.customer_id = pay.customer_id
12  WHERE city IN ('Aurora', 'Atlixco', 'Xintai', 'Adoni', 'Dhule(Dhulia)', 'Kurashiki',
13   'Pingxiang', 'Sivas', 'Celaya', 'So Leopoldo')
14  GROUP BY cu.customer_id, first_name, last_name, city, country
15  ORDER BY total_amount_paid DESC
16  LIMIT 5;
```

Data output | Messages | Notifications

	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	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