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.)**

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

   o **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.**

QUERY

SELECT D.country,

   COUNT(A.customer_id) AS customer_count

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

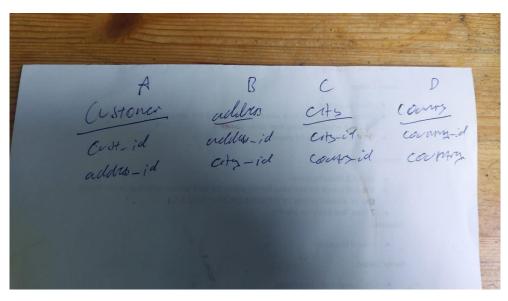INNER JOIN country D ON C.country_id = D.country_id

GROUP BY D.country

ORDER BY customer_count DESC

LIMIT 10

First I used the ERD to work out which tables and columns I needed to answer the query

Then, following the example in the lesson, started to construct the query.

I found the INNER JOINs logical and thefore quite easy.

I knew I only needed COUNT(customer_id) and country

I struggled with the syntax for COUNT(A.customer_id), initially I tried A.COUNT(customer_id). Co-pilot helped me out there
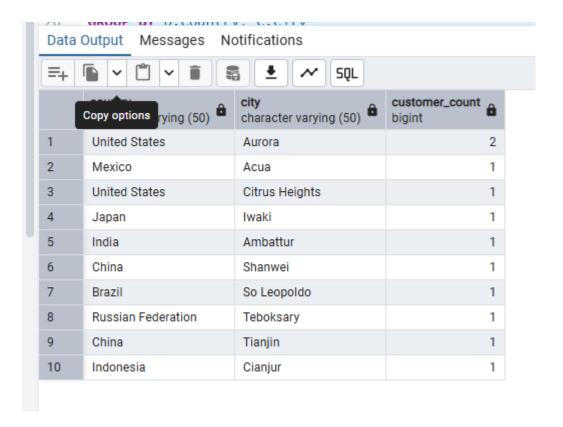
I was pleased to remember to give this an alias

I then forgot the GROUP BY Statement, but PostgresSQL suggested that fix.

I quickly had a query that worked!

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.)**

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

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

WITH top_countries AS (

  SELECT D.country

  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 D.country

  ORDER BY COUNT(A.customer_id) DESC

  LIMIT 10

)

SELECT

  D.country,

  C.city,

  COUNT(A.customer_id) AS customer_count

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 (SELECT country FROM top_countries)

GROUP BY D.country, C.city

ORDER BY customer_count DESC

LIMIT 10;

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

I couldn't get this to work. I needed Co-pilot to help me, but I now understand the process that it went through.

Essentially my whole original query is now within a WITH clause at the start.

That means that later in the query, we can call on just those Top 10 countries.

The end of the query is very similar to the original, excelpt that we also want to select C.city – and this also needs to appear withing GROUP BY

Then the key stage is to have this part of the query, so we are filtering to just the top 10 countries.

WHERE D.country IN (SELECT country FROM top_countries)

3. **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.**

My first though was to use my entire last query as a one big WITH clause (that itself contains a WITH clause.

That would work, but feels a bit complex (in case anything didn't work) and also wouldn't include a WHERE clause, which you've asked for.

SELECT A. customer_id,

A. first_name,

A.last_name,

C. city,

D. country,

SUM(F.amount) AS total_spend

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 rental E ON A.customer_id = E.customer_id

INNER JOIN payment F ON E.rental_id = F.rental_id

WHERE city = 'Aurora' OR city = 'Acua' OR city = 'Citrus Heights' OR city ='Iwaki' OR city = 'Ambattur' OR city = 'Shanwei' OR city = 'So Leopoldo' OR city = 'Teboksary' OR city = 'Tianjin' OR city = 'Cianjur'

GROUP BY A. customer_id, A. first_name, A.last_name, C. city, D. country

ORDER BY total_spend DESC

LIMIT 5;

So, I worked out which 6 tables I needed to draw data from, and constructed the JOINs for these.

I then used the WHERE to filter by the top 10 cities already established. I knew this wasn't the best way but would work

NOTE: Looked up afterwards, could have just used:
WHERE city IN ('Aurora', 'Acua', 'Citrus Heights', 'Iwaki', 'Ambattur', 'Shanwei', 'So Leopoldo',
'Teboksary', 'Tianjin', 'Cianjur')
The IN function doesn't feel intuitive to me, so I'm struggling to remember how to use it

The they key is to ORDER BY the alias for the sum of spend "total_spend"