SUBMITTED BY:
Grace Skelley on 11.5.2022

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.)**
   - *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.*

QUERY

```
SELECT D.country
      ,COUNT(customer_id) AS totalNumber_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 D.country
ORDER BY count (A.customer_id) DESC
LIMIT 10
```

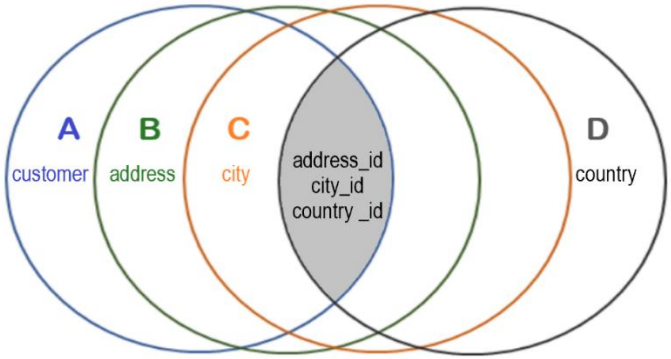

FIGURE 1: Illustration of Multiple Joins

TOP 10 COUNTRIES

Data Output | Messages | Notifications

	country character varying (50)	totalnumber_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

Total rows: 10 of 10 | Query complete 00:00:00.096

Using the ERD as part of formulating the syntax for this query, I linked each table as such by identifying first its common keys (*Fig. 1*). To present the Top 10 Countries, I began to search countries through country table D then, linking the following: address_id though A & B, city_id in B & C and country on C & D. The total number of customers is `GROUP BY` country in `ORDER BY` customer_id with total highest count.

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.*

QUERY

```
SELECT C.city AS Top10_Cities,D.country AS Top10_Country
      ,COUNT(customer_id) AS totalNumber_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 ,D.country
ORDER BY count (A.customer_id) DESC
LIMIT 10
```

TOP 10 CITIES
(WITHIN THE 10 LEADING COUNTRIES)

Data Output | Messages | Notifications

	top10_cities character varying (50)	top10_country character varying (50)	totalnumber_of_customers bigint
1	Aurora	United States	2
2	Atlixco	Mexico	1
3	Xintai	China	1
4	Adoni	India	1
5	Dhule (Dhulia)	India	1
6	Kurashiki	Japan	1
7	Pingxiang	China	1
8	Sivas	Turkey	1
9	Celaya	Mexico	1
10	So Leopoldo	Brazil	1

Total rows: 10 of 10 Query complete 00:00:00.103

With similar to the first question, I then based the script on the previous command but adding `WHERE` clause and narrowing down Top 10 Cities by executing `IN` operator retrieval within the top 10 countries. The total number of customers is `GROUP BY` city and country in `ORDER BY` customer_id with total highest count.

And polishing the results with appropriate column names for easy interpretation.

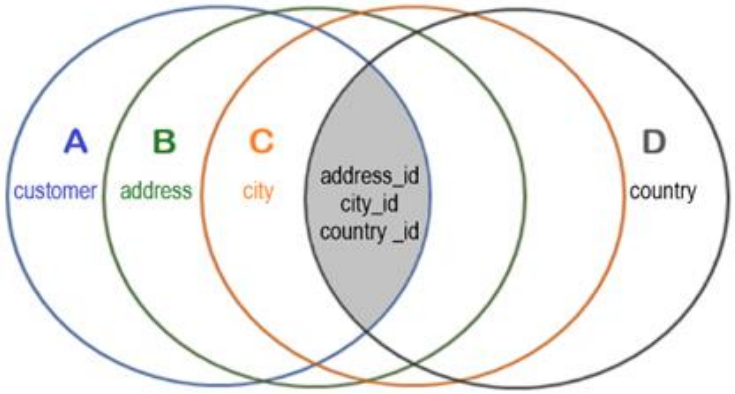

FIGURE 1: Illustration of Multiple Joins

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

```
SELECT A.customer_id
      ,A.first_name
      ,A.last_name
      ,C.city
      ,D.country
      ,SUM(E.amount) AS TotalAmount_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 E.customer_id = A.customer_id
WHERE  C.city  IN  ('Aurora'  ,'Atlixco'  ,'Xintai'
,'Adoni' ,'Dhule (Dhulia)' ,'Kurashiki' ,'Pingxiang'
,'Sivas' ,'Celaya' ,'So Leopoldo')
GROUP  BY  A.customer_id  ,A.first_name  ,A.last_name
,C.city ,D.country
ORDER BY SUM(E.amount) DESC
LIMIT 5
```

TOP 5 CUSTOMERS
(WITHIN TOP 10 CITIES)

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)	totalamount_paid numeric
1	84	Sara	Perry	Atlixco	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

Total rows: 5 of 5 | Query complete 00:00:00.120

To disclose the top 5 customers, the results must display several columns: custumer_id, first_name, last_name, city, country, total amount paid by executing `SELECT` query from customer, city, country and payment tables. I then, integrate the `INNER JOIN` syntax to include values. The relevant data points will be pulled from the top 10 cities (which was the previous search) and `GROUP BY` city and country in `ORDER BY` the highest total amount paid.

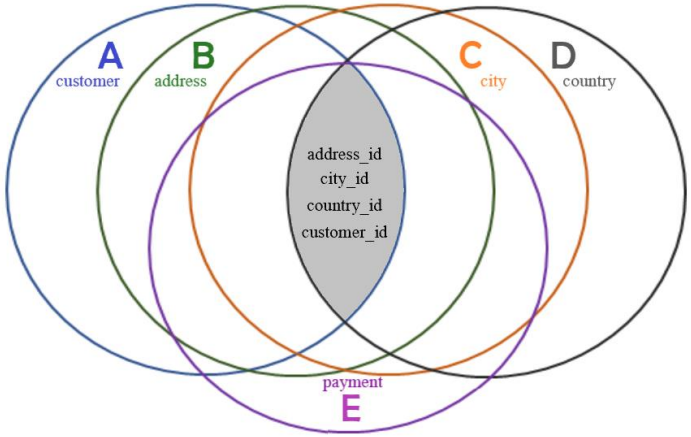

FIGURE 2: Illustration of Multiple Joins