Data Immersion Ach 03.07

Ryan Wick – 01/28/2025

Directions

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.
        - **1. Query**

Query | Query History

```
1  SELECT D.country,
2       COUNT (A.customer_id) as customers
3  FROM customer A
4  INNER JOIN address B ON A.address_id = B.address_id
5  INNER JOIN city C ON B.city_id = C.city_id
6  INNER JOIN country D ON C.country_id = D.country_id
7  GROUP BY D. country
8  ORDER BY COUNT(A.customer_id) DESC
9  LIMIT 10
10
```

        - **1. Output**

	country character varying (50)	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

- 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.
  - To start off with the problem at hand I worked to identify all the tables needed from the ERD. Once I was able to figure out all the connections between customer, address, city, and country I was ready to start writing my query. My primary focus was the customer_id and country columns. With these two in mind, I utilized the INNER JOIN function to sequentially join the tables referencing their foreign keys. Once that was working properly I added the GROUP BY, ORDER BY, and LIMIT commands to get the desired outcome.

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.
     - **2. Query**

```sql
SELECT C.city AS Top_Cities,
       D.country AS Top_Countries,
       COUNT (A.customer_id)as total_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 (
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)
GROUP BY D.country, C.city
ORDER BY COUNT(A.customer_id) DESC
LIMIT 10;
```

     - **2. Output**

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

- Write a short explanation of how you approached this query and why.
  - Since we are looking for the top 10 cities that fall within our top 10 cities I started by reusing the previous query. Because we needed to ensure that the cities are from the top 10 countries I utilized the WHERE clause to perform a subquery. From here I knew I again needed to perform an INNER JOIN to get all the relevant tables together. Next it was time to utilize the GROUP BY command to have the results grouped based off country and city. Last but not least, I used ORDER BY to get everything in descending order and LIMIT to provide only the top ten desired outputs.

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.
     - **3. Query**

Query | Query History

```sql
SELECT B.customer_id AS top_5_Customer_id,
        B.first_name AS customer_first_name,
        B.last_name AS customer_last_name,
        D.city AS top_Cities,
        E.country AS top_countries,
        SUM(A.amount) AS highest_payment
FROM payment A
INNER JOIN customer B ON A.customer_id = B.customer_id
INNER JOIN address C ON B.address_id = C.address_id
INNER JOIN city D ON C.city_id = D.city_id
INNER JOIN country E ON D.country_id = E.country_id
WHERE D.city IN
        (
        SELECT C.city
        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 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
                )
        GROUP BY C.city
        ORDER BY COUNT(A.customer_id) DESC
        LIMIT 10
        )
GROUP BY B.customer_id, B.first_name, B.last_name, D.city, E.country
ORDER BY SUM(A.amount) DESC
LIMIT 5;
```

- ### 3. Output

	top_5_customer_id integer	customer_first_name character varying (45)	customer_last_name character varying (45)	top_cities character varying (50)	top_countries character varying (50)	highest_payment numeric
1	566	Casey	Mena	Tokat	Turkey	130.68
2	84	Sara	Perry	Atlixco	Mexico	128.70
3	506	Leslie	Seward	Pontianak	Indonesia	123.72
4	389	Alan	Kahn	Emeishan	China	119.75
5	537	Clinton	Buford	Aurora	United States	98.76

4. Finally, save your “Answers 3.7” document as a PDF and upload it here for your tutor to review.