Task 3.7 – Joining Tables of Data

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

```
15  --Where top 10 customers are located
16  SELECT D.country,
17  COUNT(A.customer_id) AS customer_count
18  FROM customer A
19  INNER JOIN address B ON A.address_id = B.address_id
20  INNER JOIN city C ON B.city_id = C.city_id
21  INNER JOIN country D ON C.country_id = D.country_id
22  GROUP BY country
23  ORDER BY customer_count DESC
24  Limit 10;
```

Data Output | Explain | Messages | Notifications

	country character varying (50)	customer_count 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.

-I approached the query as an inner join because we need information from both tables to match or else, we might get null responses for countries.
-Since we need the top 10 I started with the aggregate formula of COUNT for customer_id
-Next step was to join each of the tables through their foreign or primary keys starting with address_id in the customer table all the way to country_id in the country table in order to get country (name).
-Next we needed to group country so we could get the customer count for each country.
-Then we needed to order customer_id by desc so we could limit it by the top 10 results.

2. Write a query to find the top 10 cities within the top 10 countries identified in step 1.

```sql
--Top 10 cities within the top 10 countries
SELECT C.city,
    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
WHERE country IN ('India',
                  'China',
                  'United States',
                  'Japan',
                  'Mexico',
                  'Brazil',
                  'Russian Federation',
                  'Philippines',
                  'Turkey',
                  'Indonesia')
GROUP BY country,city
ORDER BY customer_count DESC
LIMIT 10;
```

Data Output | Explain | Messages | Notifications

	city character varying (50)	country character varying (50)	customer_count 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

Similar approach to question 1 but added in the WHERE clause to only show the top 10 results from the top 10 cities found above.

3. Write a query to find the top 5 customers in the top 10 cities who have paid the highest total amounts to Rockbuster.

```sql
--top 5 customers in the top 10 cities who have paid the highest amounts
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 city IN ('Aurora','Acua','Citrus Heights','Iwaki','Ambattur','Shanwei','So Leopoldo','Teboksary','Tianjin','Cianjur')
GROUP BY A.customer_id,
    D.country,
    C.city
ORDER BY total_amount_paid DESC
LIMIT 5;
```

Data Output | Explain | Messages | Notifications

	customer_id integer	first_name character varying (45)	last_name character varying (45)	country character varying (50)	city character varying (50)	total_amount_paid numeric
1	225	Arlene	Harvey	India	Ambattur	111.76
2	424	Kyle	Spurlock	China	Shanwei	109.71
3	240	Marlene	Welch	Japan	Iwaki	106.77
4	486	Glen	Talbert	Mexico	Acua	100.77
5	537	Clinton	Buford	United States	Aurora	98.76